Point-by-point response to the reviews

Referee 2

We thank this reviewer very much for the so detailed constructive comments on this work. We have made changes to the manuscript accordingly. We colored our response in blue. Text from the manuscript is quoted with double quotation marks and new text is shown in *italics*.

In this paper, Li *et al.* investigate the sensitivity of dust in the CESM2-CAM6.1 climate model to various parameterized processes: the emissions scheme, the dry deposition scheme, the fixed geometric width of the coarse mode, and the assumption of spherical/aspherical particles. Using a wealth of validatory observations and many simulations, they find that changing dust emissions and the coarse mode width have the greatest impact on the dust metrics, followed by the dry deposition scheme and then asphericity. They also propose a new version of CAM (CAM6.α) which improves on many dust metrics relative to CAM6.1 and incorporates some of the listed process changes.

The paper is well written and contains a wealth of useful information, including the most comprehensive database of dust observations yet (Table 3). The introduction is highly readable, and the conclusions are generally supported by the analysis. However, this paper rather feels like 3 independent studies convoluted together, namely, (1) a new and improved version of the dust scheme in CAM (CAM6.1 versus CAM6.α), (2) a study of the sensitivity of simulated dust to certain processes, and (3) a study of the merits of separating dust into its mineralogical components in CAM. I think the paper would benefit from being split into 2 or 3 separate papers, which I expand on below in the General Comments. In short, I think that the study needs a redesign before it is published, which may require major revisions (i.e., new simulations and a re-write) and/or splitting into separate papers.

We appreciate the positive comments very much. The reviewer correctly pointed it out that this is a paper convoluted together by independent studies. Our original plan, however, was to separately document the size change in BULK CAM6, and the improved emission and deposition parameterizations in CAM6. The merits of modeling dust as mineral components have been very well documented before in terms of the climatic impacts by mineral dust, so we do think it would not deserve a new paper on this. Since we tend to update separate processes in CAM6 and the new schemes have been detailed and tested offline or in previous versions of CAM (CAM4 and CAM5), it makes more senses to document in the same paper how the change to each process may affect the dust cycle modeling. Please see our reply to the comments below including that regarding to the experiment redesign.

Firstly, I think that the simulation design is incorrect for exploring the sensitivity of dust to the altered processes. For example, the new dry deposition scheme is only tested in conjunction with the other altered processes (CAM6.α and CAM6.α_MINE) and never on its own. Conversely, the new emissions scheme is tested by itself for both BULK and MINE dust models, whilst the size and shape of the particles are tested in conjunction with the new emissions scheme but using BULK and MINE dust respectively. In short, it's very difficult to attribute the impacts on the dust metrics to the individual processes.

I would suggest concentrating on either the BULK dust scheme or the MINE dust scheme, unless you plan to directly compare them. The study would be much cleaner if the processes were tested in isolation using either BULK or MINE and then compared to CAM6.1 (see Table below). In its current form, it is very difficult to disentangle which dust impacts emanate from which altered process.

Suggested simulations:

Simulation	Name	Description
1	CAM6.1	Standard model
2	NEW_EMISS	CAM6.1 with BRIFT emissions
3	NEW_SIZE_S5	CAM6.1 with CAM5 size assumptions
4	NEW_SIZE_S6σ5	CAM6.1 with CAM5 assumptions except coarse σ from CAM6.1
5	NEW_DRYDEP	CAM6.1 with PZ10 dry deposition
6	NEW_SHAPE	CAM6.1 with aspherical dust
7	CAM6.1_MINE	Equivalent to MINE_BASE but may use CAM6.α as BASE simulation
8	CAM6.α	CAM6.1 with all of the relevant model changes

In summary, I would highly recommend that the authors run further simulations with each of the processes applied separately as the current simulation design is not conducive or particularly supportive of the results presented in the manuscript.

This is a similar comment to what the first Reviewer raised. Below we paste our reply to the comment by Reviewer # 1 (Line 41-169 in that report) as a response in the text.

There are a couple of different methods to estimate the effect of each development, such as the one we used and the one the reviewers suggested. Strictly speaking, either method cannot totally exclude the possible influence of the parametrizations that had already been included and can affect the dust cycle modeling in the base model, such as the advection scheme and cloud processing. The reason is that there likely exists a nonlinear "interaction" between the existing parameterizations and the newly introduced one, which seems weak though. We acknowledge that adding one by one seems clearer than the original experiment design, but it requires more simulations and thus more computational resources while yielding a similar estimate of the impact of each development (**Fig. R1**), compared to what we had presented based on our own's experiment. We had selected our own's set of experiments, because adding the modification on top of the previous change can help understand how the simulated dust cycle evolves while updating the model (MINE_BASE) toward the most advanced one (CAM6.α_MINE).

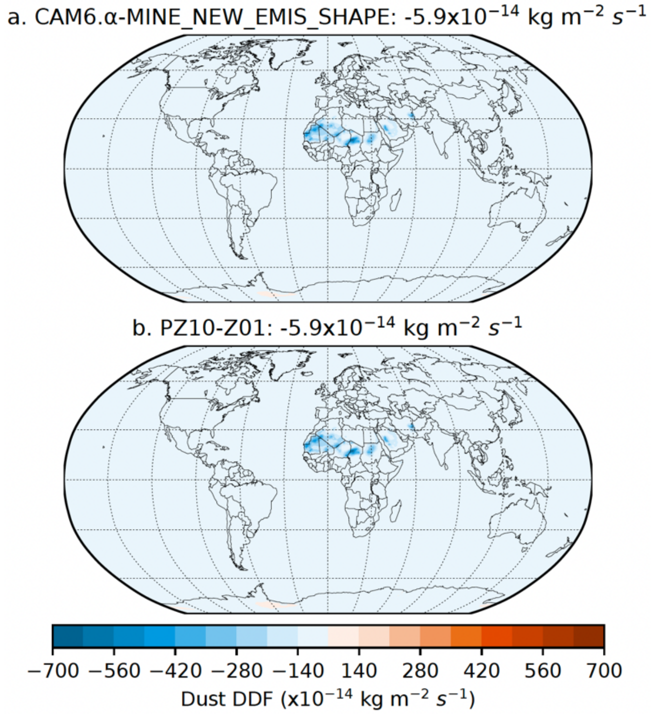


Fig. R1. Influence of changing to PZ10 on the simulated dry deposition fluxes in the dust-speciated model (change to the global annual mean of dry dust deposition: ~70 Tg) based on our experiment (a) and the suggested experiment by the reviewers (b; Simulation 5 – simulation 1). Quantified change to the global annual mean of dry dust deposition equals ~70 Tg by either method.

The BULK runs were constructed to investigate how the mistakenly set dust size distribution influences the dust cycle modeling and the estimate of dust DRE. This inappropriate size distribution has been employed in studies using the officially released BULK CAM6 and not in any study using the dust-speciated CAM. So, we do not have a good reason to perform size sensitivity tests in the MINE runs. What's more important is that quantifying the impact of individual processes, based on the base CAM6.1 that uses an inappropriate dust size distribution, seems not that meaningful: it makes more sense to make such quantification using models with the "correct" size distribution. That is why in all the MINE runs designed for that purpose we revert the narrow coarse-mode size distribution to the broad one. Also, following the reviewer's design would change little to the results obtained from our experiments on the dust cycle modeling. This is because the offline dynamics and the dust tuning employed ensures quite similar dust cycles modeled by BULK and MINE with different developments (Fig. R2 and Fig.

R3), if the size distribution is also set identical, since the sum of the mass fraction for each of the eight minerals always equals unity. We had pointed this similarity out in our originally submitted manuscript: “It is worth noting that with the dust tuning applied toward the similar global mean DOD ~0.03 the modeled dust cycle (i.e., burdens, concentrations, loadings, deposition fluxes) would be similarly comparable between the bulk and speciated dust models using the same offline dynamics and dust size distribution”. Repeating the set of simulations using BULK instead to quantify the impact of each altered process would then yield similar results that we presented in the manuscript.

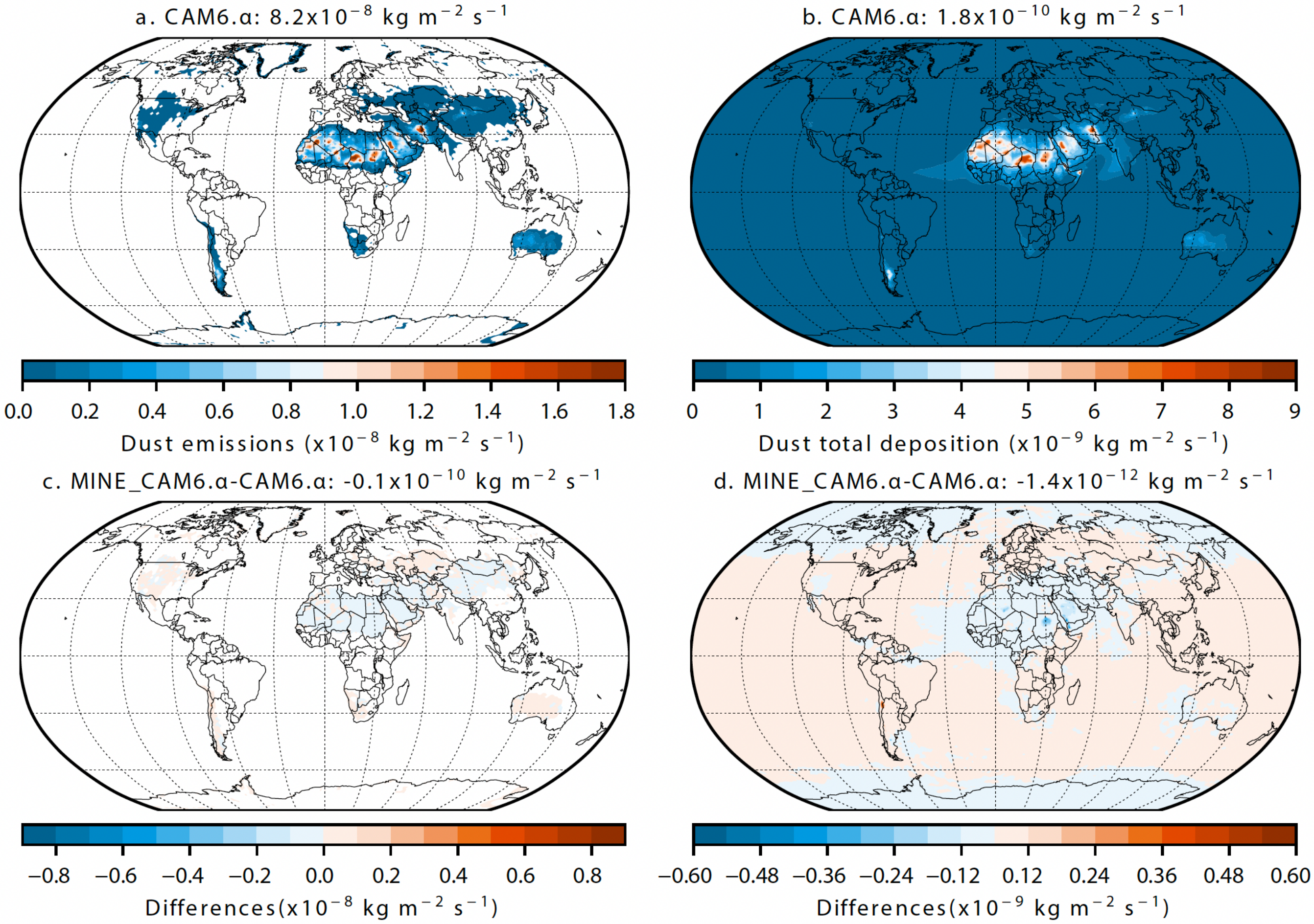


Fig. R2. Surface dust emissions (a; global annual mean=2891 Tg) and deposition fluxes (b; global annual mean=2893 Tg) simulated by CAM6.α and their differences (c and d; both global annual mean=22 Tg) between MINE_ CAM6.α and CAM6.α.

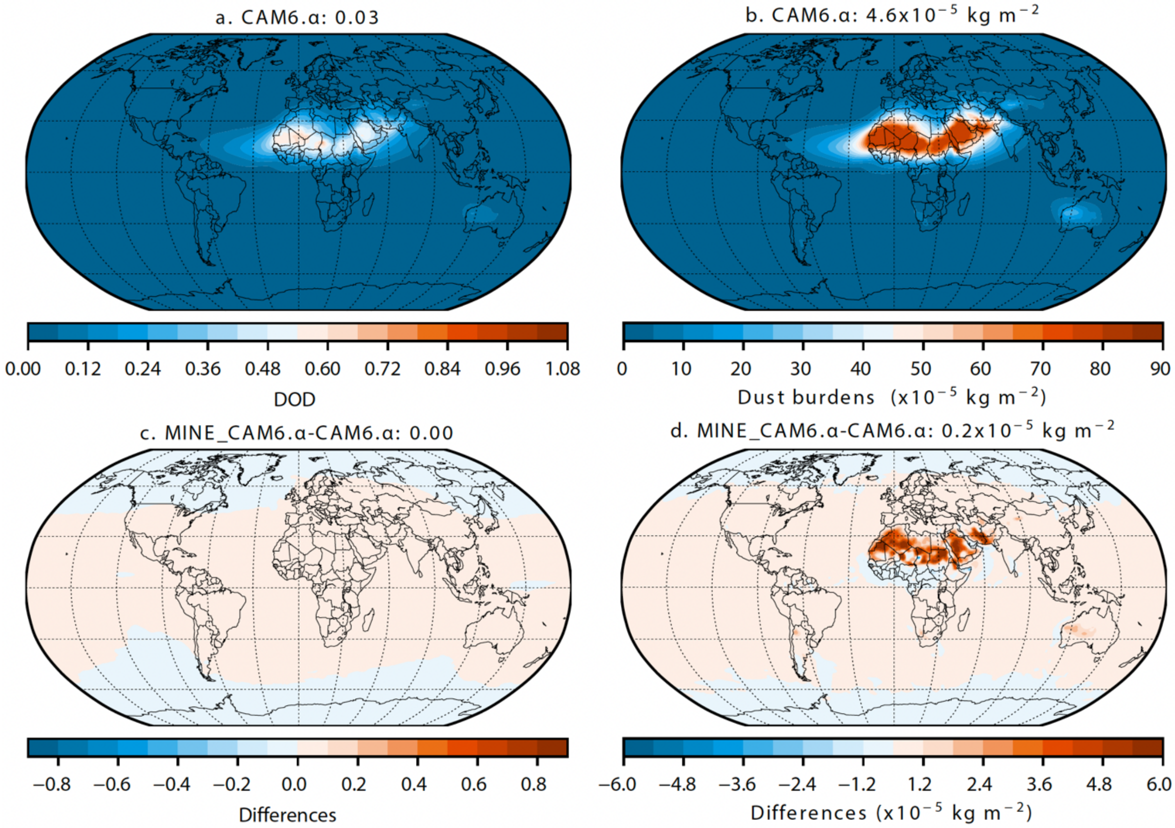


Fig. R3. The same as **Fig. R2** but for DOD (a: global annual mean=0.030 and c: global mean difference=0.001) and dust burdens (b: global annual mean of dust mass=24 Tg and d: global mean difference≈0 Tg), respectively.

NEW_EMIS and MINE_NEW_EMIS appear like a duplication of experiment for testing the new dust emission scheme. But the estimate of dust DRE differs considerably between the two experiments.

To reflect the Reviewer's suggestion, we added the following in the section "2.6 Experiment design**"**:

"*We investigate how the mistakenly set dust size distribution influences the dust cycle modeling and the estimate of dust DRE in the bulk-dust model rather the speciated-dust model, because this inappropriate size distribution has been employed in previous studies using the officially released bulk-dust CAM6 only and not in any study using the speciated-dust CAM. It is also reasonable to make all the quantifications in the model that use an appropriate dust size distribution. Therefore, we reverted the dust size distribution in all the speciated-dust runs to that configured in CAM5.*"

"*It worth noting that with the dust tuning applied toward the similar global mean DOD ~0.03 the modeled dust cycle (i.e., burdens, concentrations, loadings, and deposition fluxes) would be* *similarly comparable between the bulk- and speciated-dust models that nudged toward identical offline dynamics and using the same dust size distribution (see Sect. 6). The quantified effect of each of the modifications would thus be similar if using the bulk dust model instead (Fig. S2: R1 in this report), but the modeled dust optical properties (e.g., single scattering albedo) by the bulk and speciated dust models differ considerably, resulting in considerably different dust DRE (Scanza et al., 2015) and DRE efficiencies between NEW_EMIS (CAM6.α) and MINE_NEW_EMIS (CAM6.α_MINE).*”

“*A comparison of the BULK and MINE models on simulating dust DRE had been previously documented (Scanza et al., 2015). This study includes the MINE runs because we want to check as well if the updates help improve reproducing the observed dust DRE efficiency in a model that may more reasonably represent the regional variation of dust optical properties.* Note that there are many ways to conduct sensitivity studies, which could lead to slightly different results. *We added the modification on top of the previous change to understand how the simulated dust cycle evolves while updating the model (MINE_BASE) toward the most advanced version (CAM6.α_MINE). This may not hinder a clean comparison of the effect of each development, since the ‘interaction; between the existing and the newly introduced parameterizations appears weak (Fig. S2: R1 in this report).*”

To clarify how we quantify influence of each development, we added two columns in the Table 4 pointing out the size distribution used and purpose of each experiment and added the following text in the “Experiment design” section:

“*We separately compared the performance of PZ10 to Z01, aspherical to spherical dust, and BRIFT to DEAD on the simulated dust cycle and quantified influence of each of those modifications on the climatic-effect estimate by comparing the modeled dust cycle in the paired simulations CAM6.α _MINE vs MINE_NEW_EMIS_SHAPE, MINE_NEW_EMIS_SHAPE vs MINE_NEW_EMIS, and MINE_NEW_EMIS vs MINE_BASE, respectively.*”

Finally, we added a separate new section to compare results from BULK with those from MINE:

“*6. Bulk- versus speciated-dust model*

The bulk (CAM6.α) and dust-speciated model (CAM6.α_MINE) simulate a similar dust cycle with the difference between the two types of models orders smaller than those simulated by the former (Fig. 12 and 13: R2 and R3 in this report, respectively). This similarity results from the dust tuning toward the global mean DOD of 0.03, the same meteorology dynamics both models were nudged toward, and the design of the dust-speciated model that summing the mass fraction of each dust species equals unity. With the same reasons, the influence of each of the modifications on the modelled dust cycle quantified using the bulk model instead of the dust-

speciated model, as this study used, would be similarly comparable. The modelled dust optical properties, however, differs remarkably (i.e., dust SSA; Table 6) with the simulated global mean dust SSA by CAM6.α_MINE (0.896) lower than by CAM6.α (0.911) at the visible band centered at 0.53 μm. Note the dust DRE is sensitive to variation of the dust SSA. This lower dust SSA obtained here in the dust speciated model than in the bulk dust model is consistent to the finding of a previous study (Scanza et al., 2015) using an early model version (CAM5). Correspondingly, CAM6.α_MINE yields a reduced dust cooling (Table 6) and DRE efficiency (Fig. 9) than CAM6.α. For dust DRE efficiency (Fig. 9), speciating dust in CAM6 tends to reduce the RMSE while retaining the horizontally spatial correlation in either SW (CAM6.α: RMSE=11; R=0.26 versus CAM6.α_MINE: RMSE=10; R=0.20) or longwave (CAM6.α: RMSE=4; R=0.86 versus CAM6.α_MINE: RMSE=3; R=0.84) or both spectral ranges (CAM6.α: RMSE=7; R=0.93 versus CAM6.α_MINE: RMSE=6; R=0.92). This comparison suggests that modeling dust as component minerals with the dust size distribution in coarse mode of MINE_NEW_EMIS_SIZE helps improve the model performance relative to modeling dust as a bulk to reproduce the retrieved dust DRE efficiency (Fig. 9a). The improvement, however, could be artificial because of the combined use of imaginary complex refractive index of hematite volume (see Fig. 1b of Li et al., 2021) and the volume mixing used in the dust speciated model to compute the bulk-dust complex refractive index (Li et al. in prep.), leading to artificially more absorptive dust than in the bulk dust model (Fig. 9a and Table 6)."

Another issue that I had with the simulation design was the arbitrary tuning of dust optical depth (DOD) to 0.03 in some simulations but not in others (L289). This made it very difficult to quantify the impact of the altered processes and forced the authors to add caveats throughout the text e.g., L590 "differences between the global annual mean dust deposition in BRIFT and DEAD would become smaller, if we rescaled the value according to the same DOD criteria". I suggest only tuning CAM6.1 and CAM6.α to 0.03 and using the tuned CAM6.1 as the BASE model in which to add the different processes incrementally. I see no need to rescale DOD in the sensitivity simulations and it would be interesting to see the impact of the different processes on the global-mean DOD as a derived product of the models. Tuning to 0.03 is arbitrary and also misses the fact that much of the dust mass is in the super coarse mode which is missing from the model, and therefore the model may be wrongly tuned to 0.03.

We had tuned CAM6.1 and CAM6.α as this reviewer also suggested toward 0.030. But we respectively do not agree with the reviewer that those are the only two simulations that need the retuning. For example, we must retune the model that uses the updated dust emission scheme. This is simply because if using the same tuning parameter value as in the model with DEAD, the global mean DOD would be >15 times higher than that in DEAD, reaching up to 0.45, which is undoubtedly unrealistic. We added the following in the manuscript.

"*MINE_BASE requires the dust tuning to use a much larger tuning parameter (dust_emis_fact=3.64), compared to CAM6.1 (dust_emis_fact=0.91), because, otherwise, if using the same dust_emis_fact as in DEAD, the dust emissions in BRIFT would lead to an unrealistically high global mean DOD (>~0.5).*"

The MINE runs are not for sensitivity studies but used for quantifying how each of the modifications affects the dust cycle modeling. We would obtain the same results if performing BULK runs instead, because as stated, with the same model configurations set in this study, the BULK and MINE simulations are nearly identical in terms of modeling the dust cycle. We clearly pointed this out in the revised manuscript as below.

"*It is worth noting that with the dust tuning applied toward the similar global mean DOD ~0.030 the modeled dust cycle (i.e., burdens, concentrations, loadings, and deposition fluxes) would be similarly comparable between the bulk and speciated dust models using identical offline dynamics and dust size distribution. The quantified effect of each of the modifications would be thus similar if using the bulk dust model instead.*"

As to the dust mass distribution with respect to dust size, according to a recent study (Di Biagio et al., 2020), for a total 39 Tg dust, approximately 33% (13) Tg dust are particles >10 μm, though such estimates were obtained based on model simulations. However, this missing fraction of "super-coarse" dust constitutes only a small fraction of the total DOD <2% which is even much smaller than the uncertainty in the best estimate from Ridley et al. (2016). Therefore, we believe missing that dust mass would not affect the accuracy of tuning dust toward DOD ~0.030.

In response to the reviewer's question about the dust tuning, we added some words to very briefly explain why and how we tuned the model to get the global mean DOD ~0.030.

"*We prefer to tuning the model to reproduce the global mean DOD, 0.030, because DOD is currently the best estimate of global dust quantities, compared to the others (i.e., dust concentrations). It turns out that doing so can also reasonably reproduce the other quantities with no need of a regional tuning. We tuned the dust model by modifying a namelist variable in CAM, called soil_erod_factor, corresponding to λ in Eq. (16).*"

Regarding the reviewer's suggestion to include the updates one by one, please see our response to the previous comment on the experiment design (Line 64-159).

It is also confusing for the reader that some simulations have emissions scaled by 1/f_clay whilst others have the scaling as 1, and so the impact of this change is difficult to disentangle using the current suite of simulations. It would be better if this factor is consistent across the simulations or tested in isolation.

We thank the reviewer for the comment which makes us realize that our writing may be confusing. This inversed clay fraction for the tuning factor b is not used in any of those simulations to calculate the threshold gravimetric water content. To improve the readability and to not rely excessively on external references, we introduced more the parameterization for both emission schemes (please see Section 2.5 in the revised manuscript) and added a new column in Table 2 showing the b value used in each of those experiments.

"**Table 2.** Simulations performed in this study for years 2006-2011. Treatment of dust tracer: speciated dust with separate tracers (MINE: mineralogy), or no dust speciation (bulk); the dust emission scheme: Zender et al., (2003a; DEAD) or Kok et al., (2014a; BRIFT); with or without accounting for the lifetime effect of dust asphericity (Asp versus Sph); dry deposition scheme: Zhang et al., (2001; Z01) or Petroff and Zhang (2010; PZ10); parameters for size distribution taken from the released version of CAM5 and CAM6.1 (see Table 1 for CAM5 and CAM6 size, respectively); additional test on dust size distribution using the coarse-mode σ=1.2 from the released version of CAM6.1 and the rest parameters (e.g., boundaries of the geometric mean diameter) from the released version of CAM5; meteorology field nudged toward reanalysis data (offline) for 2000s climate; *dust tuning parameter includes the CAM namelist variable (dust_emis_fact) and b used in the calculation of the threshold gravimetric water content (see Sect. 2.5.1). The variable* f_{clay} *denotes the clay fraction in CLM5.* CAM6.1 and CAM6.α in bold refer to the default model and proposed new model versions, respectively, with bulk dust. Note negligible influence on the dust cycle modeling and corresponding DRE by changing the size parameters of the accumulation mode between CAM5 and CAM6 size."

Exp.	Case names	Dust model	Dry dep.	Lifetime effect of dust asphericity	Emi. scheme	*Dust size distribution*	*Dust tuning parameters (dust_emis_fact; b)*	Comments
01	**CAM6.1**	Bulk	Z01	No (Sph)	Zender [2003a]	*Default CAM6 size (Table 1)*	$0.91; 100*f_{clay}$	Officially released version
02	NEW_EMIS	Bulk	Z01	No (Sph)	Kok [2014a]	*Default CAM6 size (Table 1)*	$28; 100*f_{clay}$	Control for size tests
03	NEW_EMIS_SIZE	Bulk	Z01	No (Sph)	Kok [2014a]	*Default CAM5 size (Table 1)*	$28; 100*f_{clay}$	Changing the coarse-mode size distribution; influence quantified by comparing this with Exp. 02
04	NEW_EMIS_SIZE_WIDTH	Bulk	Z01	No (Sph)	Kok [2014a]	*Default CAM6 size but with width of the coarse-mode size distribution from defaulted CAM5 size*	$28; 100*f_{clay}$	No change to size parameters for the other modes; influence quantified by comparing this with Exp. 02
05	**CAM6.α**	Bulk	PZ10	Yes (Asp)	Kok [2014]	*Default CAM5 size*	*3.6; 1.0*	New bulk dust model
06	MINE_BASE	Mine	Z01	No (Sph)	Zender [2003a]	*Default CAM5 size*	$1.6; 100*f_{clay}$	Baseline for quantifying the impact of each modification
07	MINE_NEW_EMIS	Mine	Z01	No (Sph)	Kok [2014a]	*Default CAM5 size*	*3.6; 1.0*	Changing the dust emission scheme: influence quantified by comparing this with Exp. 06
08	MINE_NEW_EMIS_SHAPE	Mine	Z01	Yes (Asp)	Kok [2014a]	*Default CAM5 size*	*3.6; 1.0*	Experiment for changing the dust emission and shape

09	CAM6.α _MINE	Mine	PZ10	Yes (Asp)	Kok [2014a]	*Default CAM5 size*	*3.6; 1.0*	New mineralogy dust model: combined influence of the new emission scheme, PZ10, and dust asphericity quantified by comparing this with Exp. 02

Below is part of the new text relevant to the tuning factor.

“*Because of the neglection of the non-erodible elements, u*t is mostly determined by soil moisture content, which means that the augmentation factor of u*t is:*

$$f_{*t} = \begin{cases} \sqrt{1 + 1.21(w - w')^{0.68}}; w > w' \\ 1; w \leq w' \end{cases} \quad (8)$$

Where w and w’ are soil moisture content and the threshold gravimetric water content of the top soil layer.

Fécan et al. (1999) parameterized the threshold gravimetric water content (w) of the top soil layer by

$$w' = b\left(17 f_{clay} + 14 f_{clay}^{2}\right), (9)$$

where w is in percentage and b is a tuning factor.

Equations (8) and (9) are also used in DEAD with an equivalent tuning factor b set to be 1/f_{clay} *which in BRIFT is set as unity. The clay fraction is taken from the FAO(2012) soil database (see Fig. S1 of Kok et al., 2014).*”

Note b is set to be 100*f_{clay} as part of the DEAD emission scheme used in the default CAM6 but is set to be unity in BRIFT to well reproduce the observations:

“*An offline sensitivity test (Table S1: R1 in this report) supports the use of unity tuning factor to calculate the threshold gravimetric water content which we employed in the experiments for quantifying influence of each modification (speciated dust simulations listed in Table 2).*”

“Table S1 (R1 in this report). Comparison of the three CESM simulations with the offline dynamics and different values of the tuning parameter (b) to calculate the threshold gravimetric water content in the new dust emission scheme, against measurements. The measurements include AERONET AOD climatology, surface dust concentrations, and dust deposition fluxes, as described in Section 3.”

Parameter b	Correlation coefficient (RMSE) on climatology		
	AERONET DOD	Surface dust concentrations (log space)	Dust deposition fluxes (log space)
0.5	0.74 (0.13)	0.83 (0.66)	0.72 (0.93)
1.0	0.68 (0.14)	0.82 (0.72)	0.77 (0.86)
2.0	0.66 (0.14)	0.83 (0.66)	0.79 (0.82)

I gather from the text (L649) that the impact of asphericity on the dust mass extinction efficiency (MEE) is represented in *all* of these simulations. This is rather confusing, as it suggests some representation of asphericity is incorporated even when dust is assumed to be spherical (?). Please clarify this for the reader. In particular, please state whether the impact of asphericity on MEE is only applied in the simulation with dust asphericity or in all simulations (which seems inherently wrong). Really these details should be included in the Methods (L98, L224) and not in the result section.

Such impact of the dust asphericity is included in all the simulations because we did not attempt to quantify such effect in this study. To avoid of confusion and clarify this, we moved relevant text from the result Section 4.2.3 (5.2.3 in the revised version) to Section 2.4.3, and added the following to the "Experiment design" section:

"*The enhancement of the mass extinction efficiency of aerosol particles by dust asphericity is included in all the simulations, since we do not attempt to quantify how this enhancement impacts the simulated dust cycle, which has been previously well documented (Kok et al., 2021).*"

In terms of the presentation of the results, I thought that comparing CAM6.α _with CAM6.1 before looking at the individual processes was confusing, as much of the analysis of the impacts of individual processes could have been used to explain differences between the dust metrics in CAM6.1 and CAM6.α.

We think either doing what the reviewer suggested or keeping as what it was should work. In the drafted manuscript, we had tried doing the same as the reviewer suggested but then reordered the result section taking the "principle" that "the most important goes first", since the manuscript is lengthy. In any order, the conclusions of this article would remain unchanged.

Additionally, the authors say the following in Section 2.5:

“It is worth noting that dust burdens and deposition fluxes would be comparable, if the bulk and speciated dust models have similar DOD. But the dust optical properties (e.g., single scattering albedo) in the bulk and speciated dust simulations differ, resulting in considerably different dust direct radiative effects and direct radiative effect efficiencies. Therefore, we state the difference in the dust DRE and DRE efficiency estimate in Sect. 6, but do not document the comparison of dust loadings/deposition/DOD between the bulk and speciated dust simulations.”

Given that DOD is tuned to be similar in these simulations, I do not see why the differences in optical properties should be used as an excuse not to compare BULK with MINE. This would be a very interesting study in its own right, and possibly the authors should omit MINE simulations in this paper as without comparing BULK with MINE, it is difficult to understand why MINE is used at all. Is the additional mineralogical detail in MINE useful for a better dust simulation? What is the additional computational expense of MINE over BULK? Is MINE being considered for inclusion in a future of CAM or is this rather an interesting pedagogical study? Currently, MINE is frivolously used in this study and is unnecessary without further analysis and comparison.

We did not compare the modeled dust cycle between BULK and MINE runs, because this is science with secondary importance. We show the reason in the “Experiment design” section: “It is worth noting that with the dust tuning applied toward the similar global mean DOD ~0.03 the modeled dust cycle (i.e., burdens, concentrations, loadings, and deposition fluxes) would be similarly comparable between the bulk and speciated dust models using identical offline dynamics and dust size distribution. The quantified effect of each of the modifications would be thus similar if using the bulk dust model instead.”

The different optical properties are not the reason for not making the comparison but are one of the reasons for why we included the MINE runs: we have shown in the text evaluations on the model performance of modeling the DRE efficiency and the influence of each modification on the DRE estimate, for which modeling the optical properties as accurately as possible is crucial. Therefore, dust speciated model is better to use to quantify such influence, as it simulates spatially varying dust optical properties, while the bulk dust model is using a globally constant dust optic.

We had tried to do this but found that having the potential impacts on the mineralogy by changing to the new dust emission scheme is not enough for a separate paper. Instead, we added more analysis on documenting results from the MINE runs, such that it makes more senses to have both BULK and MINE runs in this article.

“***6. Bulk versus speciated-dust model***

The bulk (CAM6.α) and dust-speciated model (CAM6.α_MINE) simulate a similar dust cycle with the difference between the two types of models orders smaller than those simulated by the former (Fig. 12 and 13). This similarity results from the dust tuning toward the global mean DOD of 0.03, the same meteorology dynamics both models were nudged toward, and the design of the dust-speciated model that summing the mass fraction of each dust species equals unity. With the same reasons, the influence of each of the modifications on the modelled dust cycle quantified using the bulk model instead of the dust-speciated model, as this study used, would be similarly comparable. The modelled dust optical properties, however, differs remarkably (i.e., dust SSA; Table 6) with the simulated global mean dust SSA by CAM6.α_MINE (0.896) lower than by CAM6.α (0.911) at the visible band centered at 0.53 μm. Note the dust DRE is sensitive to variation of the dust SSA. This lower dust SSA obtained here in the dust speciated model than in the bulk dust model is consistent to the finding of a previous study (Scanza et al., 2015) using an early model version (CAM5). Correspondingly, CAM6.α_MINE yields a reduced dust cooling (Table 6) and DRE efficiency (Fig. 9) than CAM6.α. For dust DRE efficiency (Fig. 9), speciating dust in CAM6 tends to reduce the RMSE while retaining the horizontally spatial correlation in either SW (CAM6.α: RMSE=11; R=0.26 versus CAM6.α_MINE: RMSE=10; R=0.20) or LW (CAM6.α: RMSE=4; R=0.86 versus CAM6.α_MINE: RMSE=3; R=0.84) or both spectral ranges (CAM6.α: RMSE=7; R=0.93 versus CAM6.α_MINE: RMSE=6; R=0.92). This comparison suggests that modeling dust as component minerals with the dust size distribution in coarse mode of MINE_NEW_EMIS_SIZE helps improve the model performance relative to modeling dust as a bulk to reproduce the retrieved dust DRE efficiency (Fig. 9a). The improvement, however, could be artificial because of the combined use of imaginary complex refractive index of hematite volume (see Fig. 1b of Li et al., 2021) and the volume mixing used in the dust speciated model to compute the bulk-dust complex refractive index (Li et al. in prep.), leading to artificially more absorptive dust than in the bulk dust model (Fig. 9a and Table 6).”

Specific comments

[L75] Is it worth introducing the DEAD and BRIFT acronyms here?

Introduced here.

[L84] The fine mode is described as d < 1um whilst the coarse mode is d > 5um. Normally, the coarse mode is adjacent to the fine mode so I wonder what the authors would define the intermediate aerosol (1 < d< 5um) as?

We just follow the definition normally used in the community. So, here is not a definition for the coarse mode aerosol. To avoid of possible confusions, we revised this statement as below:

“*…and slightly underestimating that of aerosols with diameter > 5.0μm…*”.

[L91] *"one of the changes from CAM5 to CAM6.1 was replacing the size distribution of aerosols in the coarse mode in CAM5 with the one that has a much narrower width in CAM6.1"*- this seems nonsensical to me, or completely without consideration for actual coarse mode dust widths (e.g., Ryder et al, 2013, 2018, 2019 suggest $\sigma \in [1.6, 2]$ rather than 1.2). Why was it decided to favour stratospheric sulfate over tropospheric mineral dust when sulfate is more episodic (e.g. volcanic eruptions) and has less of an impact over tropospheric climate? Also, the authors seem to recommend that the coarse mode width be reverted to 1.8 as in CAM5 (I agree), but do not comment on the impact of resetting the coarse mode width on stratospheric sulfate. Seeing as this was the initial motivation for contracting σ, I think that some comment is appropriate.

That is right: we also think 1.2 is too narrow to use to represent size distribution of dust aerosol, so we had decided to revert it to 1.8 in this work with which this reviewer also agree and recommend using this broad size distribution in the future versions of CAM. In CAM6, the volcanic sulfate is presented together with dust aerosol. The developers were focusing on the volcanic sulfate while advancing the CAM model without noticing that the employed sigma is inappropriate for dust aerosol.

We commented a little bit on this as below.

"*Our analysis suggests that the defaulted 1.2 for the geometric standard deviation of the transported dust size distribution (coarse mode) may be too narrow to simulate the dust lifetime.* In the next released model version, we recommend reverting the geometric standard deviation to 1.8, as in CAM5, *which may require a split of representation of dust and the stratospheric aerosols. It is this reversion that imposes the most important change among what we introduced to CAM6.1 to the modeled dust cycle.*"

[Table 1] I think that GMD should be labelled as "initialisation GMD" _as this is more descriptive. Or is the initial GMD at source calculated online? It is difficult to tell from the text what the initial GMS is. This also refers to L179.

Changed to "initialized GMD" here and elsewhere it is applicable. The reviewer is right that this is initialized GMD.

[Table 1] Why is the order of the modes Accumulation, Aitken, Coarse, then Primary? Surely it should be in ascending size order: Primary, Aitken, Accumulation then Coarse

The order in the table 1 is the same as that in the model which lists the accumulation mode ahead of the Aitken mode and primary mode at the last. In response, we reordered the list following the reviewer's suggestion.

"**Table 1.** Mode parameters for the Modal Aerosol Module version 4 (MAM4) used in CAM5 (CAM5 size) and CAM6.1 (CAM6 size) by default: geometric standard deviations (σ) and initialized geometric mean diameter (GMD) and its ranges. Values in parentheses if present are for CAM6.1 cells without parentheses are kept the same between CAM5 and CAM6.1."

Mode (note order)	σ	*Initialized* GMD (μm)	Lower bound GMD (μm)	Upper bound GMD (μm)
Primary carbon (a4)	1.6	0.050	0.010	0.10
Aitken (a2)	1.6	0.026	0.0087	0.052
Accumulation (a1)	1.8(1.6)	0.11	0.054	0.44
Coarse (a3)	1.8(1.2)	2.0(0.90)	1.0(0.40)	4.0(40)

[Table 1] Why was the accumulation mode width changed in CAM6.1? What are the impacts of reverting it? I can't see this detail in the text

Good point. It is the same reason for this slight change as that in the coarse mode to accommodate the stratospheric aerosol (Mills et al., 2016), but our test simulations suggest negligible impacts on the dust cycle modeling when reverting it. We very briefly mentioned this in the revised manuscript (see the "Experiment design" section).

"*The other changes to the width of the accumulation mode and the bounds of the simulated GMD online impose negligible impacts on the dust cycle modeling, thus, we did not construct sensitivity tests on reverting them in this study.*"

[L109] The term 'semi-observation' is undefined and is confusing

We now specify both the observation and semi-observation as "measurements, retrievals, and model-observation integration" which brackets all the data used in this work.

[L115] "show the final summarization in Section 7". This is an unusual way to say "Discussion and conclusions are provided in Section 7" or something to that effect

We changed it to:

"*…limitations in the model-observation comparison in Sect. 5, and discussions and conclusions in Sect. 7.*"

[L120] This is one of the places in the text where it is unclear as to: (1) whether the impact of dust asphericity on MEE is represented at all, (2) if it is represented then in what way (methods), and (3) which simulations include it?

To avoid of confusion and clarify this, we removed "and optics" here, moved relevant text from the result Section 4.2.3 (5.2.3 in the revised manuscript) to Section 2.4.3, and added the following to the "Experiment design" section:

"*The enhancement of the mass extinction efficiency of aerosol particles by dust asphericity is included in all the simulations, since we do not attempt to quantify how this enhancement impacts the simulated dust cycle, which has been previously well documented (Kok et al., 2021)*."

[L137] Sentence beginning *"We consider the default DEAD scheme"* should explicitly acknowledge that it refers to emissions

Changed "scheme" to "dust emission scheme".

[L143] How confident are the authors in the critical LAI threshold? Should this assumption be discussed in the Discussion section?

The relationship of the bare soil fraction and LAI and the critical LAI threshold has been used as a standard for a while in CAM of different versions. It could be subject to change in the future, but the associated uncertainty would probably be smaller than that due to what we discussed in the Discussion section which are important missing pieces for modeling dust aerosols in CAM6. Still, we added one sentence in response to this good question.

"*This large uncertainty could partially result from the constants used in the parametrizations that affect the dust emission and transport processes, such as the critical LAI threshold, though it has been used during the past decade in different CAM versions.*"

[L152] The mass is distributed as 0.1 %, 1 % and 98.9 % between the Aitken, accumulation, and coarse modes. Surely these ratios should change depending on the assumed coarse mode width?

These values were obtained by applying the brittle fragmentation theory to the broad coarse-mode size distribution, so, they are applicable to the proposed new models. But the default CAM6.1 is using the same values while employing a much narrower coarse-mode size distribution, which could be problematic.

[L160] Many dust schemes treat dust as initially insoluble and then permitted to age via coagulation and condensation wherein it becomes soluble and internally mixed (e.g., dust in UKESM1). The authors should comment on their assumption of internally mixing dust, which may artificially enhance dust deposition near source regions? Would you expect similar results if dust is assumed to be insoluble?

The internal mixing assumption within each mode has been employed as an option in CAM since the version 5 and has been made in a huge number of studies, using CAM particularly. It worth pointing it out that dust aerosols are not completely internally mixed in MAM4 of CAM5/6: dust in different modes are externally mixed. But most dust mass is distributed in the coarse mode, which indicates that the assumption to the coarse-mode dust would be most influential on the dust cycle modeling, compared to that to the other modes. In this paper, we do not mean to document how different mixing assumptions affect the dust modeling in CAM6, since all our simulations stick to this assumption. So, we only try to briefly answer the question of this reviewer but will not expand it in the manuscript. From the view of the dust cycle modeling, we think the importance of dust hygroscopicity and its mixing with other aerosols is regionally dependent. For example, a different assumption of mixing with sea salt for South African dust can greatly change simulated deposition near the source and particularly in the downwind area. But for North African dust, it is not important near the source because both cloud fractions and sea salt concentrations are typically low. But from the view of modeling the optical properties and radiative effects of dust, the mixing states really matters.

[L165] The Neale et al (2010) reference is an internal document, which I can't find online. Can the authors please provide a URL for downloading the report, or alternatively, relevant peer-reviewed papers with the same information.

RESPONSE: It's a technical note. We put it in GitHub and a link in the manuscript where we cite this reference: https://github.com/L3atm/LLi2022GMD.

[L172] *"The wet deposition rate thus depends on the hygroscopicity of dust (=0.068; Scanza et al., 2015) as CCN/INPs and the prescribed scavenging coefficient (=0.1; Neale et al., 2010), both of which are currently constant with respect to the dust size (and composition for speciated dust) in CAM6.1."* _– _I assume the hygroscopicity of dust will evolve as dust is transported through the atmosphere so I question the use of a single spatially uniform constant for this parameter. The below cloud scavenging coefficient (0.1), if it is in units of s-1, seems 2 orders of magnitude too high (Wang et al., 2010, doi:10.5194/acp-10-5685-2010). Wang et al (2010) for instance, suggest it's somewhere between 10-6 for accumulation mode aerosol and 10-3 for coarse mode aerosol depending on scavenging rate. The authors should comment more on the assumptions made in the model and the implications of those assumptions.

We appreciate the great comment and agree that the dust hygroscopicity would vary from one region to another and change during transport due to the dust ageing. But since the purpose of this paper is to document the changes, we made to CAM6.1, and how they work in effect to the dust cycle modeling, we tend to not spend too much text on commenting on all the parameterizations, such as the oversimplified hygroscopicity of dust in CAM6.1. Still, this is a very useful comment, as it could change the wet deposition rate. So, we very briefly pointed this out in the discussion section:

"*This large uncertainty could probably in part result from the constants used in the parametrizations that affect the dust emission and transport processes, such as the critical LAI threshold, the hygroscopicity of dust, and the prescribed scavenging coefficient, though the default values in the model has been used during the past decade in CAM of different versions.*"

[L180] *"Note that the current default CAM6.1 employs a narrow coarse-mode size distribution but a broad boundary width (high bound minus low bound), likely resulting in the GMD bounds less in effect, compared to that in CAM5".* – _what are the impacts of changing the coarse mode width on sea-salt emissions and sea-salt AOD? Surely this change will impact more than dust alone, which may be confounding other results presented in the study (e.g., the DRE).

This change does affect the emissions and optical depth of sea salt. We had included such impacts but then removed relevant text, since the focus of this study is on dust aerosol. Documenting sea salt seems somewhat distract the readers. To reflect this suggestion, now we mention sea salt a little bit in the last section.

"*This reverting may require a split of representation of dust and the stratospheric aerosols in the coarse mode, for which the narrow coarse-mode size distribution works better (Mills et al., 2016), and some changes to sea salt.*"

[L210] "The wet size due to growth of aerosol particles by adsorbing water vapor follows the κ-Kohler theory with a time-invariant hygroscopicity for each aerosol species (Petters and Kreidenwei, 2007)". – is it worth listing these hygroscopicity parameters to aid in the replicability of the simulations?

We will archive the model code which contains the values used for each of the aerosol species and is publicly available.

[L215] "here and hereafter unless stated otherwise" – this phrase, in parentheses, doesn't seem to apply to anything or make sense

Removed.

[L224] This is another place in the text where the impact of asphericity on the MEE is tantalisingly hinted at without further detail as to whether its on and how its incorporated

We clarified this in Section 2.5 of the revised manuscript as below, so, here removed "calculated mass extinction efficiency and".

"*The enhancement of the mass extinction efficiency of aerosol particles by dust asphericity is included in all the simulations, since we do not attempt to quantify how this enhancement impacts the simulated dust cycle, which has been previously well documented (Kok et al., 2021)*"

[L276] "In addition, the meteorology field (horizontal wind, air temperature T, and relative humidity) was nudged" – the results will obviously be changed if the model is free running then. For instance, the coarse dust will absorb LW radiation, warming the surface and destabilising the atmosphere. Perhaps this assumption (fixed meteorology) should be discussed in the Discussion section

The reviewer is right. If a free running is constructed, which we will do in the future, the results could be different. We had pointed it out that the results here are from simulations based on the use of offline dynamics in the first paragraph of the last section. To emphasis this, at some other places in the Discussion section, we inserted "offline dynamics":

"*It worth noting that the results obtained in this study rely on the models with the offline dynamics, which is subject to change while using the predicted meteorology field online.*"

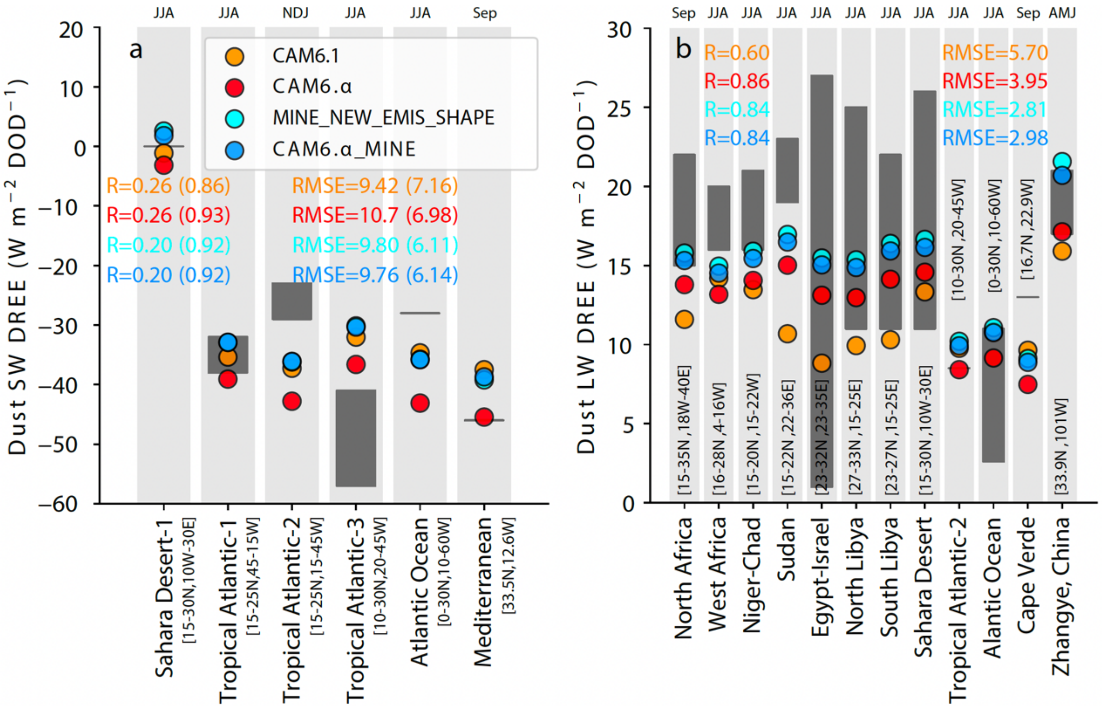
"*…with the offline dynamics, the new model, CAM6.α…*"

[L285] *"Therefore, we state the difference in the dust DRE and DRE efficiency estimate in Sect. 6, but do not document the comparison of dust loadings/deposition/DOD between the bulk and speciated dust simulations."* – Avoiding comparing BULK and MINE seems like a massive oversight and is one of the first things I'd query as a reader. Does speciation between minerals improve the simulation compared to assuming dust as a bulk quantity? Simply saying that as the dust properties are different (of course they will be), this reduces comparability, is a little bit absurd and a bit of a cop out. I think this comparison should be made in a follow-on paper. To be honest, it doesn't seem worth including the MINE simulations if they not appropriately analysed.

The dust speciation helps better reproduce the observed DRE efficiency improvements compared to without the speciation, as presented in the Section 4.3.1 (5.3.1 in the revised text). For non-optical variable, summing over the eight minerals gives the total dust loadings/deposition/DOD similarly comparable to that from simulations without the dust speciation. Per the suggestion of the reviewers, we added a new section "*6. Bulk- versus speciated-dust model*" collecting information about the comparison between BULK and MINE results that scattered in the text: s

"*This lower dust SSA obtained here in the dust speciated model than in the bulk dust model is consistent to the finding of a previous study (Scanza et al., 2015) using an early model version (CAM5). Correspondingly, CAM6.α_MINE yields a reduced dust cooling (Table 6) and DRE efficiency (Fig. 9) than CAM6.α. For dust DRE efficiency (Fig. 9), speciating dust in CAM6 tends to reduce the RMSE while retaining the horizontally spatial correlation in either SW (CAM6.α: RMSE=11; R=0.26 versus CAM6.α_MINE: RMSE=10; R=0.20) or LW (CAM6.α: RMSE=4; R=0.86 versus CAM6.α_MINE: RMSE=3; R=0.84) or both spectral ranges (CAM6.α: RMSE=7; R=0.93 versus CAM6.α_MINE: RMSE=6; R=0.92). This comparison suggests that modeling dust as component minerals with the dust size distribution in coarse mode of MINE_NEW_EMIS_SIZE helps improve the model performance relative to modeling dust as a bulk to reproduce the retrieved dust DRE efficiency (Fig. 9a). The improvement, however, could be artificial because of the combined use of imaginary complex refractive index of hematite volume (see Fig. 1b of Li et al., 2021) and the volume mixing used in the dust speciated model to compute the bulk-dust complex refractive index (Li et al. in prep.), leading to artificially more absorptive dust than in the bulk dust model (Fig. 9a and Table 6).*"

Also, we added RMSE and correlation coefficient in the DRE efficiency plot as shown below.



“**Figure 8.** Modelled and observed dust direct radiative effect efficiency in the shortwave (SW)/longwave (LW) spectral ranges under clear conditions at the TOA over the sub-domains (shown in the inserted map and location described below) *in April-June (AMJ), summer (JJA), fall (NDJ), and September (Sep)* for the 2000s climate. The radiative effect efficiency is defined as the ratio of the radiative effect to DOD, so has units of W m^{-2} τ^{-1}. Included cases from left are CAM6.1, CAM6.α, MINE_NEW_EMIS_SHAPE, CAM6.α _MINE. The field value/range are from references listed in Table 5. *Colored numbers show correlation coefficient (R) and the root mean square error (RMSE) between the model and retrievals in the SW (a) / LW (b) spectral ranges or in both spectral ranges (numbers in parenthesis in Panel a).*”

[L289] Choosing to tune some models to DOD = 0.03 but not others is very peculiar. The authors say *“Dust tuning was not applied to EXP03 and EXP04 (bulk dust simulations), in which the dust emission was identical to EXP02, in order to see how changes in the transported dust size distribution affects the DOD calculation”.* – Well surely all of the individual sensitivity simulations (emissions, dry deposition, asphericity) would have benefitted from the same analysis? I guess that some parameters in the emissions and dry deposition algorithm need to tuned in some way (so using DOD might be a reasonable approach) as the parameters have a huge degree of uncertainty, but the asphericity probably did not need changing.

Though we did not tune EXP03 and EXZP04 which we had previously run, we scaled up DOD and applied the same factor to the other dust quantities, as we stated in the text (the “Experiment design” section). This rescaling makes senses, considering the roughly linear relationship between those variables, though we acknowledge uncertainty may introduce by doing so. We pointed this out in the Discussion section.

“…*though the linear assumption between DOD and the other dust quantities based on which we rescaled up the concentrations, deposition, burdens, and DRE of dust in the size distribution simulations*.”

In the emission and deposition schemes, we agree that there could maybe exist large uncertainty in some parameters. But we would better not scale the non-tunable parameters within the dust scheme to match the observational constraint of DOD=0.03, because the scaling factor exists largely due to the missing sub-grid scale variability by 100-km grid-scale modeling, not because of the uncertainty of parameters. Tuning those parameters to match the global constraint just seems like errors compensating each other. The dust emission scheme in CAM contains a tuning parameter “b”, in the calculation of the threshold gravimetric water content, which can plausibly range from less than 1 to the inversed clay fraction (can be > 3.0). Sensitivity tests by modifying this tuning parameter among 0.5, 1.0, and 2.0 suggest that 1.0 is a good value to use (see Table R1 in this report). We would not change parameters that are not introduced as tunable ones, since they are observationally constrained. It is for this reason we did not modify parameters in the new dry deposition scheme considering that those are all non-tunable. We added the following in the “Experiment design” section and cited a new supplementary table (Table R1) there:

“*An offline sensitivity test (Table S1) supports the use of unity tuning factor to calculate the threshold gravimetric water content which we employed in the experiments for quantifying influence of each modification (speciated dust simulations listed in Table 2).*”

[L289] My other issue with this paragraph is that the tuning is not described in any detail. Which parameters were tuned and what are their values in the baseline simulation? How was tuning conducted and why was global-mean DOD chosen as the target? Simply saying ‘tuned the model following Albani et al (2014)’ _is not sufficient, and it would be impossible to replicate these simulations without further detail

We added the following to address this comment.

“*…we tuned the model following Albani et al., (2014) by modifying a namelist variable called soil_erod_factor, such that…*”

“*We prefer to tuning the model to reproduce the global mean DOD, 0.030, because this is currently the best estimate of global dust quantities, compared to the others (i.e., dust concentrations). It turns out that doing so can also reasonably reproduce the other quantities with no need of a regional tuning. We tuned the dust model by modifying a namelist variable in CAM, called soil_erod_factor.*”

[Table 3] This table seems very large, and I’m not sure whether the list of acronyms should be at then end of the table or in the caption. Would it be better to have 1 table for each metric?

We split this large table into 3 and list the acronyms in the caption:

“***Table** 3. Observed/retrieved cycle for dust model evaluations including optical depth, surface mass concentrations, surface deposition fluxes, and wet deposition percentages. AERONET: Aerosol Robotic Network; MODIS: Moderate Resolution Imaging Spectroradiometer; AOD: aerosol optical depth; DOD: dust optical depth.*”

“***Table** 4. Measured/retrieved dust size distribution for model evaluation. AERONET: Aerosol Robotic Network; DustCOMM: Dust Constraints from joint Observational-Modelling-experiMental analysis.*”

*“**Table** 5. Retrieved dust radiative effect efficiency for model evaluation. CERES: Clouds and the Earth’s Radiant Energy System; TOA: top of the atmosphere; JJA: June, July, and August; AOD: aerosol optical depth; MISR: Multi-angle Imaging SpectroRadiometer; OMI: Ozone Monitoring Instrument; NDJ: November, December, and January; MODIS: Moderate Resolution Imaging Spectroradiometer; CALIPSO: Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations; MFRSR: MultiFilter Rotating Shadowband Radiometer; SEVIRI: Spinning Enhanced Visible and Infrared Imager; GERB: Geostationary Earth Radiation Budget; AERONET: Aerosol Robotic Network; MPL: Micro-Pulse Lidar; AERI: Atmospheric Emitted Radiance Interferometer; SMART: Surface-sensing Measurements for Atmospheric Radiative Transfer; AMJ: April, May, and June.”*

[Results] The difference between CAM6.1 and CAM6.α i.e., the control and the simulation with all changes added (except mineralogy) comes before the dissection of impacts of individual processes. Why is this? Surely it would be better to investigate the impacts of the individual processes and then use them to explain why CAM6. α is different to CAM6.1?

We think either doing what the Reviewer #2 suggested or keeping as what it was should work. In the drafted manuscript, we had tried doing the same as the reviewer suggested but then reordered the result section taking the “principle” that “the most important goes first”. But in any order the conclusions of this article remain unchanged.

[L378] *“CAM6.1 may overestimate the contribution of high-latitude dust emissions to the global dust total (8.0%).”* – is this referring to the dust burden? It’s rather ambiguous as is

This refers to the dust emission. We modified this sentence a little bit.

“CAM6.1 may overestimate the contribution of the high-latitude dust emission to the global dust total *emission* (8.0%).”

[L391] *“Overall, all models reproduced the climatology of DOD from AERONET retrievals, the surface concentration, and deposition within a factor of ten (Fig. 1 and Fig. S3)”* – this doesn’t seem to be the case from looking at Fig. 1 b, c, e, f, h, and i. It seems that both models exhibit at least one measurement outside the range of 1/10x and 10x.

There are only 1, 4, and <10 point(s) of the 36, 47, and 108 points for DOD, surface concentrations, and deposition outside that range. That is, for over 90% of the points fall in the factor of 10. To be more accurate, we modified the sentence a little bit as follows.

“Over 90% of the measurement sites, all models reproduced the climatology of DOD from AERONET retrievals, the surface concentration, and deposition within a factor of ten (Fig. 1 and Fig. S3)”

[Fig. 2] Why is the new dust emissions scheme smoother in terms of emissions, rather than the delta function (almost) in DEAD? I couldn’t easily find this information in the text

We added the following to answer this question.

“*The smoother distribution of the dust emission in BRIFT than DEAD is due primarily to the use of the source function in DEAD that shifts dust emissions toward the most erodible soil, while in BRIFT, the near-surface friction velocity frequently exceeds the calculated threshold wind fraction velocity, which seems low in the land model, causing dust to emit at more grid cells.*”

[Fig. 3] Isn’t the Ridley et al (2016) DOD dataset constrained by MODIS (either through assimilation or using it as a baseline? If so, aren’t Figs 3a and 3b effectively showing the same results?

Good point. DOD of Ridley et al. (2006) “assimilated” MODIS retrievals: they corrected the bias present in MODIS retrievals (see Section 3 in the manuscript), so the former contains information of the latter, but the two datasets show considerably different results. For example, the globally averaged DOD from pure MODIS postprocessed by Pu et al. (2020) is significantly higher than the best estimate of Ridley et al. (2016) (0.025-0.035).

[L436] capture -> captures

Corrected.

[L437] Taklamakan (as in the desert) is spelt wrong throughout

Corrected.

[Fig. 4] Great figure

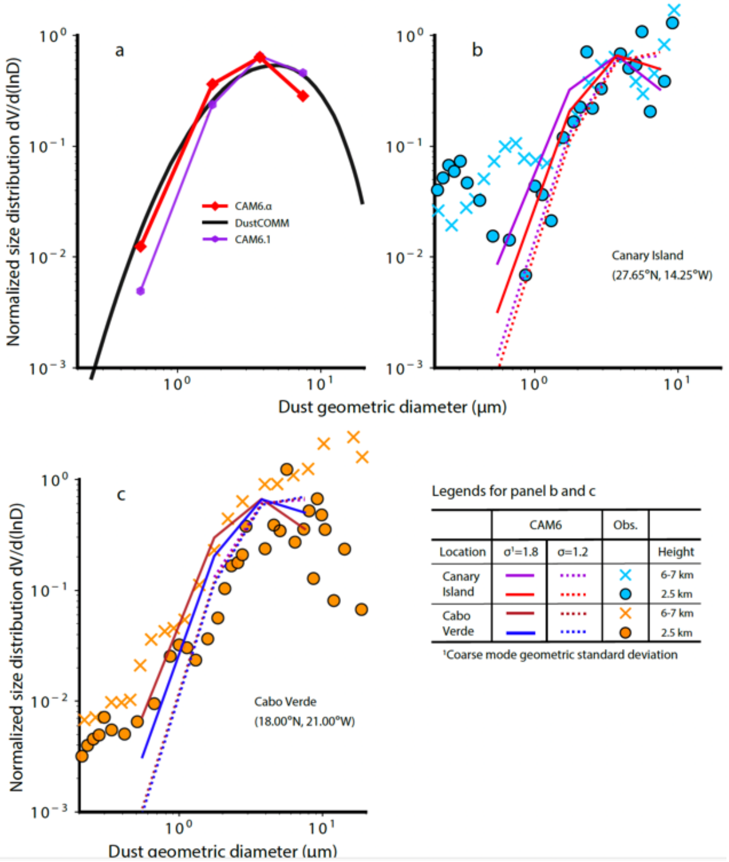
Thanks!

[L498] S5i -> S5e

Corrected.

[Fig. 5] This plot, especially Fig. 5a, is very confusing. There are too many colours and it is difficult to pick out the CAM models. It may be worth plotting a non-CAM multi-model mean with max/min as shaded in grey, and then have just the CAM models in colour

We removed non-CAM model results and cited relevant reference instead.



"**Figure 5.** Normalized size distribution of dust between 0.2 and 10 μm diameter in the global average (a), near Canary Island (blue colors in b; dot: 2.5 km; x: 6-7 km; data for June/July 1997 from Otto et al., 2007), and near Cabo Verde (orange colors in c; dot: 2.5 km; x: 6-7 km; data for August 2015 taken from Ryder et al., 2018). The default model, CAM6.1: (purple line); the new model, CAM6.α: (red line); semi-observations: DustCOMM (black line) *inverted based on an integration of a global model ensemble and quality-controlled observational constrains on the transported dust size distribution, extinction efficiency, and regional DOD* with data taken from Adebiyi et al. (2020). We chose the model layers and grid cells that are closest to the location and atmospheric height, as well as the months, where and when the measurements were made for comparison."

[L542] Why is the size distribution for the fine dust fraction better captured by CAM6.α?

We explained this in the revised manuscript.

"*CAM6.α can better capture due primarily to the more accurate gravitational settling velocity modeled by using the new dry deposition scheme.*"

[L548] Sentence beginning "Overall, CAM6.α better reproduced the size distribution". It would be worth adding the caveat here that the Otto et al and Ryder et al measurements are from single campaigns or flights and thus may not reflect the long-term mean dust properties at those altitudes, locations, and times

Good point. We had introduced a separate section listing out limitations that are commonly presented in the model-data comparison which includes this point. But it's good to mention again at this place. So, we added the following.

"*It is worth noting that the measurements are from single campaigns or flights that may have representative issues not reflecting the climatological size and vertical distributions of dust aerosols (i.e., limited by the space and time coverage).*"

[L558] Section 4.2.1 – why are the mineralogy experiments used to test BRIFT vs DEAD rather than the BULK simulations? There doesn't appear to be any reasoning behind this

BULK and MINE runs were originally designed in two separate papers, but we ended up with this one. The results would be very similar between using BULK and MINE to test BRIFT vs DEAD and the other schemes. See our response to the general comment on BULK versus MINE (Line 61-190).

[L559] MIINE_NEW_EMIS -> MINE_NEW_EMIS

Done.

[L646] Paragraph on asphericity – I'm still confused even after reading the text as to whether the assumption of asphericity is applied to the dust MEE in every simulation run here or just the MINE_NEW_EMIS_SHAPE simulation?

We paste our response to the previous comments of this reviewer here:

Such impact of the dust asphericity is included in all the simulations because we did not attempt to quantify such effect in this study. To avoid of confusion and clarify this, we removed "and optics" here, moved relevant text from the result Section 4.2.3 (5.2.3 in the revised manuscript) to Section 2.4.3 (2.5.3 in the revised manuscript), and added the following to the "Experiment design" section:

“*The enhancement of the mass extinction efficiency of aerosol particles by dust asphericity is included in all the simulations, since we do not attempt to quantify how this enhancement impacts the simulated dust cycle, which has been previously well documented (Kok et al., 2021).*”

[L683] “(0.030-0.019)/0.030*100)” – I don’t think this formula needs to be written. See also L686 and L759

Deleted.

[L693] Paragraph beginning *“The lifetime of dust”*. Should this paragraph be in Section 4.2.4? It doesn’t seem to mention asphericity or apply to the MINE_NEW_EMIS_SHAPE simulation

We had included this paragraph here, because this is the section that we talked about the dust lifetime: one of the main impacts of the dust size change is on the dust lifetime. Sine this is a comparison between BRIFT and DEAD, we moved to Section 5.2.1 (revised manuscript) “Dust emission schemes: BRIFT versus DEAD”.

[L705] Why is MINE_NEW_EMIS referred to as the reference case? It’s a sensitivity simulation, isn’t it? Surely the only reference cases are CAM6.1 and possibly MINE_BASE?

It is not a sensitivity simulation. The coarse-mode size distribution of dust in CAM6.1 was wrongly put. Thus, it seems not make a lot of senses to use CAM6.1 as the baseline simulation when quantifying the impact of each of the modifications. Please see our response to the general comment on the experiment design (Line 61-190).

[L733] *“NEW_EMIS_SIZE”* -> MINE_NEW_EMIS_SIZE. Also, this paragraph seems to be the only place where BULK and MINE are explicitly compared. I think the comparison should extend to all the dust metrics

As stated in our responses to previous comments, with the dust tuning and offline dynamics applied, speciating dust does not yield considerably different dust quantities (i.e., dust concentrations, burdens, and deposition) from BULK runs in the current climate. We added a new section to compare BULK and MINE runs:

“**6. Bulk- versus speciated-dust model**

The bulk (CAM6.α) and dust-speciated model (CAM6.α_MINE) simulate a similar dust cycle with the difference between the two types of models orders smaller than those simulated by the former (Fig. 12 and 13: R2 and R3 in this report, respectively). This similarity results from the *dust tuning toward the global mean DOD of 0.03, the same meteorology dynamics both models were nudged toward, and the design of the dust-speciated model that summing the mass fraction of each dust species equals unity. With the same reasons, the influence of each of the modifications on the modelled dust cycle quantified using the bulk model instead of the dust-speciated model, as this study used, would be similarly comparable. The modelled dust optical properties, however, differs remarkably (i.e., dust SSA; Table 6) with the simulated global mean dust SSA by CAM6.α_MINE (0.896) lower than by CAM6.α (0.911) at the visible band centered at 0.53 μm. Note the dust DRE is sensitive to variation of the dust SSA. This lower dust SSA obtained here in the dust speciated model than in the bulk dust model is consistent to the finding of a previous study (Scanza et al., 2015) using an early model version (CAM5). Correspondingly, CAM6.α_MINE yields a reduced dust cooling (Table 6) and DRE efficiency (Fig. 9) than CAM6.α. For dust DRE efficiency (Fig. 9), speciating dust in CAM6 tends to reduce the RMSE while retaining the horizontally spatial correlation in either SW (CAM6.α: RMSE=11; R=0.26 versus CAM6.α_MINE: RMSE=10; R=0.20) or LW (CAM6.α: RMSE=4; R=0.86 versus CAM6.α_MINE: RMSE=3; R=0.84) or both spectral ranges (CAM6.α: RMSE=7; R=0.93 versus CAM6.α_MINE: RMSE=6; R=0.92). This comparison suggests that modeling dust as component minerals with the dust size distribution in coarse mode of MINE_NEW_EMIS_SIZE helps improve the model performance relative to modeling dust as a bulk to reproduce the retrieved dust DRE efficiency (Fig. 9a). The improvement, however, could be artificial because of the combined use of imaginary complex refractive index of hematite volume (see Fig. 1b of Li et al., 2021) and the volume mixing used in the dust speciated model to compute the bulk-dust complex refractive index (Li et al. in prep.), leading to artificially more absorptive dust than in the bulk dust model (Fig. 9a and Table 6).*"

[L798] "Overall, replacing the size distribution of dust aerosol and the dust emission scheme with new ones (PZ10 and BRIFT, respectively)" – replacing the size distribution is referred to here as PZ10 but this is the dry deposition scheme

"(PZ10 and BRIFT, respectively)" removed.

[L821] The term "space volume" is ambiguous. Possibly "colocation in space"?

Changed.

[L833] *"which can get mixed with dust aerosol particles during the transport and may not be completely excluded in the measurements."* _This seems a little lazy, do you have any estimates of how much contaminations leads to errors in measuring dust? At the moment, this point isn't backed up by evidence.

We deleted these texts, because 1) the second half of the sentence reads more like a repeat of the first half which the references we had cited serve well to support, and 2) they do not convey vital elements (we compiled the dust measurements from previous publication and use them here to evaluate the model performance. This section discussed the "Limitation in the model-observation comparison", noticing the readers that such kind of error exists in the dust measurements would be fine). Sentence now reads:

"*Finally, the modelled dust mass is for dust with our own defined mineralogy composition only (Li et al., 2021; Scanza et al., 2015), the measured mass could likely also include non-dust particles, such as sea salt (Kandler et al., 2011; Zhang et al., 2006), sulfate (Kandler et al., 2007), biomass burning aerosols (Ansmann et al., 2011; Johnson et al., 2008), or other air pollution aerosol (Huang et al., 2010; Yuan et al., 2008).*"

[L859] "… followed by the enhanced dust mass extinction efficiency at the visible band by ~30% to account for the enhancement by dust asphericity" – the asphericity applied to the MEE has not been shown to be the second most important change affected. Rather Fig. 10 shows that asphericity has a negligible impact on dust. Or is the asphericity in the MEE applied separately to the asphericity in the deposition rate? This is very confusing.

We do not plan to estimate the optical effect of the dust asphericity. That is why we include such effect in all the simulations. In response, we removed text relevant to the optical effect of the dust asphericity in the conclusion section. We also clarified how we dealt with the enhanced dust mass extinction efficiency at the visible band in the simulations in the "Experiment design" section:

"*The enhancement of the mass extinction efficiency of aerosol particles by dust asphericity is included in all the simulations, since we do not attempt to quantify how this enhancement impacts the simulated dust cycle.***"**

[L869] "Overall, the new model can:" – is the new model, referred to in this sentence, CAM6.α? If so, why has CAM6.α_MINE been neglected? The addition of MINE to this study makes little sense as it is peripheral. Additionally, is this "new model" _already adopted for the next revision of CAM6 or is this the plan for the future?

We specified the new model. As to the modeled dust cycle, CAM6.α_MINE and CAM6.α show almost identical results. Please see our response to the general comment (Line 61-190). The modifications made to CAM6.1 to get CAM6.α is on the table. But the dust speciation is not planned yet to be included in a future CAM version.