The manuscript presents a comprehensive analysis of dust effective radiative forcing (DuERF) using CMIP6 ESMs under the AerChemMIP piClim-2xdust experiment, contrasted with the piClim-control run. The study decomposes DuERF into direct and indirect components, evaluates model uncertainties, and quantifies the impact of dust perturbations on precipitation. While the manuscript is well-structured and provides valuable insights, several key issues need addressing to improve clarity, interpretability, and robustness. Below, I outline major revisions the authors may want to consider to ensure the study meets its full potential.

Reply: *We thank the reviewer for the valuable comments and suggestions, this has really helped elevate the quality of the manuscript. Our reply to the reviewer comments are written in italic and line changes are referring to the tracked changes version of the revised manuscript. Blue text refers to added text.*

General comments

Comment: 1. The introduction is quite lengthy and detailed, and it gives me the impression that the authors aim to constrain the uncertainty in dust forcing estimates by addressing the factors that contribute to this uncertainty. However, this is not the primary goal of their study which I did not realize until the last several sentences of the introduction section. To better align with the study's goals the introduction could be streamlined by shortening the discussion of factors contributing to uncertainty and focusing more on how those factors are represented in ESMs and how they influence the DuERF.

Reply: *We thank the reviewer for their insightful feedback regarding the introduction. The introduction has been revised accordingly, please see the tracked changes version of the revised manuscript for the specific changes.*

Comment: 2. The study analyzes DuERF using the idealized piClim-2xdust experiment, which perturbs dust emissions under preindustrial conditions. While the emission increase (70-105%) is comparable to or larger than estimated historical changes (Kok et al., 2023), the experimental design, fixed SSTs and no anthropogenic forcing, means the results reflect model-specific dust responses rather than real-world historical forcing. The authors should clarify that these findings cannot be directly compared to studies quantifying DuERF over the industrial era, as the mechanisms and climate feedback differ. This distinction is critical to avoid misinterpretation of the DuERF values presented.

Reply: *We thank the reviewer for their valuable comment regarding the interpretation of our results in the context of the piClim-2xdust experiment. We have made changes to the manuscript accordingly, and we now explicitly state the distinctions between the DuERF derived from the piClim-2xdust experiment and model simulations targeted to quantify the real world DuERF.*

Line 171 - 178: We define DuERF as the difference in the TOA imbalance between *piClim-2xdust* and *piClim-control*, with the dust emission perturbation being the only factor that separates the two simulations.

Although the relative increase in dust in the *piClim-2xdust* is comparable in magnitude to the estimated real world historical change, it is important to note the distinction between DuERF and dust effects diagnosed from this idealised setting and real-world historical dust forcing. Specifically, sea surface temperatures (SSTs) are fixed, anthropogenic aerosols are set to preindustrial conditions, and the change in dust emission is imposed uniformly across dust source regions. Therefore, our findings cannot be directly compared with studies quantifying DuERF during the historical era (Leung et al., 2025). However, this idealised setting is still useful for investigating how ESMs behave in response to changes in dust burden.

Comment: 3. The authors need to explicitly state whether the ESMs account for LW aerosol scattering, if any, both in the model physics and in the reported LW and NET (SW+LW) forcing estimates. Omitting LW scattering likely biases the NET DuERF, as it can contribute 50% (20-60%) of the dust LW direct radiative effect at the TOA (Dufresne et al., 2002). If LW scattering is not included, please consider estimating the potential bias in the LW and NET DuERF and, if possible, adding this information to the LW and NET forcing estimates.

Also, while the manuscript notes weak LW direct forcing efficiencies (Figure 2b), it does not clarify whether this is due to missing physics (e.g., neglect of LW aerosol scattering, not only for coarse and super-coarse dust particles, although this is likely the major source), deficiencies in the size distribution (e.g., underrepresentation of super-coarse dust), or a combined effect of the two. I request clarification on these aspects.

Reply: *We thank the reviewer for highlighting these important points. All of the ESMs considered in our study account for the LW absorption of dust. However, only UKESM1-0-LL includes both LW absorption and scattering. The ESMs with a large mass extinction coefficient (MEC) are likely exhibiting deficiencies in the size distribution of dust particles. For instance, NorESM2-LM and EC-Earth3-AerChem have MEC values that align more closely with PM2.5 dust rather than PM20 (Kok et al., 2021), which indicates underestimation of coarse dust.*

To address these issues, we have updated the table to include a column indicating whether each model represents LW scattering. We have also added a discussion as a part of the conclusion section of the text regarding the impact of missing LW scattering on DuERF. ***See line 650 - 665****.*

~~The~~ Differences in the model size distribution of dust particles are an important cause of spread in simulated LW direct DuERF. ~~reflects model differences in the dust particle size distribution.~~ Despite several models claiming that they use a more realistic size distribution ~~according to~~ at the point of emissions following brittle fragmentation theory (BFT) (Kok, 2011)~~for the dust emission process~~, the large variability in dust ~~burden~~ load (larger than ~~that of~~ dust AOD) indicates a high variability in coarse dust ~~loading~~ load between ESMs. Models that include a greater fraction of coarse to coarse-super dust ~~to coarse dust do have the highest dust burden and show the largest positive LW direct DuERF efficiency per AOD. This is consistent with previous studies showing, that, after observationally constraining the size distribution to larger particles, dust exhibits less cooling (e.g., Kok et al., 2017)~~can exhibit LW forcing efficiency that is orders of magnitude larger than models that underrepresent coarse and super-coarse dust. Furthermore, ~~increasing super-coarse to coarse dust fractions would in reality cause substantial LW~~ even ESMs that include dust size distribution more aligned with observational constraints would likely still underestimate LW direct DuERF due to neglecting LW scattering, which was only included in one of the nine AerChemMIP ESMs. Including LW scattering could increase direct LW DuERF by 20% to 60% ~~of the net TOA forcing (Dufresne et al., 2002). However, this seems to be largely neglected anyway in the AerChemMIP models given the generally weak LW forcing efficiency that we find~~(Dufresne et al., 2002).

Comment: 4. The outlier behavior of NorESM2-LM is attributed to a known bug (line 400) that largely disables heterogeneous ice nucleation in mixed-phase clouds (lines 171-172). However, it is unclear how the authors handle the indirect forcing results from this model, particularly in relation to the indirect forcing. Was this model excluded from the analysis or given reduced weight in the ensemble mean? The authors should either quantify the impact of this bug on the results or provide a clear justification for including the model and explain how its known bias was accounted for.

Reply: *We thank the reviewer for their observations regarding NorESM2-LM. This model was included in the multimodel mean for the indirect forcing estimates, and we weighted all models equally that provided the necessary diagnostics for assessing cloud DuERF.*

NorESM2-LM was included in the multimodel mean for the indirect forcing estimates. We weighted all the models that provided the diagnostics needed for diagnosing the cloud DuERF equally. The bug in NorESM2-LM made the behavior of NorESM2-LM more similar to the models that did not have aerosol-aware INP representation for mixed-phase clouds. Therefore we do not judge NorESM2-LM to be any more unreliable than the models that do not represent the impact of dust INP at all. In contrast to most of the models, NorESM2-LM actually still takes dust effects on cirrus clouds into account, making it relatively advanced when it comes to dust influence on cold clouds despite the bug that disabled any dust influence on mixed-phase clouds.

To clarify, we have added a detailed explanation of this bug in the methods section (see also response to Reviewer 1, Comment 1).

5. The manuscript provides a good analysis of dust-cloud interactions and precipitation responses. However, the model descriptions in Table 1, and, in fact, throughout the manuscript, lack critical details about how cloud microphysics and precipitation processes are represented in each ESM. Only limited information is scattered throughout the text. Given that these processes are central to the study's conclusions, particularly regarding dust indirect forcings and hydrological cycle impacts, a more comprehensive summary of model physics would significantly strengthen the interpretation of results.

Reply: *First and foremost, we do not examine the microphysics in the models, while acknowledging that the models have quite different microphysical schemes. Please note that the CMIP6 models and their schemes have been extensively documented in other sources. Our main point is that the precipitation response we see in the models is more of a thermodynamic response than dust acting to drastically alter precipitation through different microphysics. We use the ARC, which is a bulk thermodynamic quantity, and that has been shown to explain reductions in* *precipitation (e.g., Pendergrass and Hartmann (2014)). We also point to the influence of absorption by dust to drive part of the inter-model differences in ARC. We acknowledge that dust also might impact ARC differently in the models due to differences in microphysics, but we deem further investigation of such impacts out of scope for this study, mainly because of missing diagnostics. We have added, however, some relevant microphysics information regarding representation of dust as INP and CCN to Table 1.*

Pendergrass, A. G., & Hartmann, D. L. (2014). The Atmospheric Energy Constraint on Global-Mean Precipitation Change. https://doi.org/10.1175/JCLI-D-13-00163.1

6. The manuscript would benefit from greater consistency in its use of aerosol optical metrics to characterize dust impacts. While the study appropriately focuses on dust radiative forcing, there appears to be an inconsistent application of optical depth metrics throughout the analysis. At times, the authors examine total aerosol optical depth (AOD) and absorption aerosol optical depth (AAOD), while at other points they reference dust optical depth (DOD). Unless I missed it, the authors did not provide dust absorption optical depth (DAOD). It is also possible that the authors are using AAOD to refer to what is defined here as DAOD. However, this should be clarified to avoid confusion.

This variation in metrics could potentially lead to confusion when interpreting results, as total AOD/AAOD incorporates contributions from all aerosol species, not just dust. Given that the study specifically investigates dust impacts through the piClim-2xdust experiment, the consistent use of dust-specific metrics (DOD and DAOD) would provide a more direct and unambiguous estimate of dust aerosol effects.

Reply: *We thank the reviewer for this insightful comment regarding the consistency of aerosol optical metrics in our manuscript. The model diagnostic that we use is the difference in AOD between piClim-2xdust and piClim-control, and since dust is the only factor (apart from minor aerosol responses in a dust-perturbed preindustrial atmosphere) impacting the change in AOD it does correspond to DOD. For the absorption AOD, note that DAOD is not a standard AerChemMIP diagnostic and thus the only way to get the dust absorption component is by taking the difference. We have updated the text according to the reviewer's suggestion. To avoid confusion with AOD, we have decided, as the reviewer suggested, to refer to the change in AOD and AAOD as DOD and DAOD throughout the manuscript.*

7. The references are not properly cited. I have pointed this out in a few places, but I encourage the authors to check for similar issues elsewhere in the manuscript.

I notice that the authors super frequently cite some frequencies, which are undoubtedly very valuable contributions to the research community. However, it is important to also acknowledge and appropriately credit the original studies that underpin key findings.

Reply: *We agree with the reviewer, it is important to value and acknowledge the original publications and have done our best to improve on this aspect the manuscript*

Minor comments

Comment: Line 35-36: What does this "environment changes" specifically refer to?

Reply**:** *This line has been removed in the revised introduction.*

Comment: Line 51-52: This statement is not strictly accurate. Several ESMs are capable of representing dust aerosols as distinct mineral components. A few examples are provided below, although the list is not exhaustive and continues to grow as model development advances.

This list here (and after) is provided for my convenience, as the items are more readily available to me than the others. The authors should feel entirely free to cite any of them or any other sources they find relevant:)

Gómez Maqueo Anaya, S., Althausen, D., Faust, M., Baars, H., Heinold, B., Hofer, J., Tegen, I., Ansmann, A., Engelmann, R., Skupin, A., Heese, B., and Schepanski, K.: The implementation of dust mineralogy in COSMO5.05-MUSCAT, Geosci. Model Dev., 17, 1271–1295, https://doi.org/10.5194/gmd-17-1271-2024, 2024.

Li, L., Mahowald, N. M., Miller, R. L., Pérez García-Pando, C., Klose, M., Hamilton, D. S., Gonçalves Ageitos, M., Ginoux, P., Balkanski, Y., Green, R. O., Kalashnikova, O., Kok, J. F., Obiso, V., Paynter, D., and Thompson, D. R.: Quantifying the range of the dust direct radiative effect due to source mineralogy uncertainty, Atmos. Chem. Phys., 21, 3973–4005, https://doi.org/10.5194/acp-21-3973-2021, 2021.

Gonçalves Ageitos, M., Obiso, V., Miller, R. L., Jorba, O., Klose, M., Dawson, M., Balkanski, Y., Perlwitz, J., Basart, S., Di Tomaso, E., Escribano, J., Macchia, F., Montané, G., Mahowald, N. M., Green, R. O., Thompson, D. R., and Pérez García-Pando, C.: Modeling dust mineralogical composition: sensitivity to soil mineralogy atlases and their expected climate impacts, Atmos. Chem. Phys., 23, 8623–8657, https://doi.org/10.5194/acp-23-8623-2023, 2023.

Li, L., Mahowald, N. M., Kok, J. F., Liu, X., Wu, M., Leung, D. M., Hamilton, D. S., Emmons, L. K., Huang, Y., Sexton, N., Meng, J., and Wan, J.: Importance of different parameterization changes for the updated dust cycle modeling in the Community Atmosphere Model (version 6.1), Geosci. Model Dev., 15, 8181–8219, https://doi.org/10.5194/gmd-15-8181-2022, 2022.

Perlwitz, J. P., Pérez García-Pando, C., and Miller, R. L.: Predicting the mineral composition of dust aerosols – Part 1: Representing key processes, Atmos. Chem. Phys., 15, 11593–11627, https://doi.org/10.5194/acp-15-11593-2015, 2015.

Scanza, R. A., Mahowald, N., Ghan, S., Zender, C. S., Kok, J. F., Liu, X., Zhang, Y., and Albani, S.: Modeling dust as component minerals in the Community Atmosphere Model: development of framework and impact on radiative forcing, Atmos. Chem. Phys., 15, 537–561, https://doi.org/10.5194/acp-15-537-2015, 2015.

Reply*: *We thank the reviewer for the additional references. The reviewer is right that many models do have the ability to represent the distinct mineral components of dust, but we have yet to find any examples of models using this representation for their CMIP production simulations due to the large computational burden including additional tracers for each mineral component. However, the original sentence did not properly reflect these nuances, and we have now rewritten the text to better reflect those nuances.* ***See line 77-90

Comment: Line 54: The current citation refers to observational data; however, the authors should instead cite modeling studies that have incorporated updated refractive indices consistent with these measurements. Two relevant studies are listed below. The first directly incorporates the measured refractive indices, while the second constrains iron oxide optics to ensure that the modeled dust optical characteristics are consistent with the measurements.

Wang, H., Liu, X., Wu, C., Lin, G., Dai, T., Goto, D., ... & Shi, G. (2024). Larger dust cooling effect estimated from regionally dependent refractive indices. Geophysical Research Letters, 51(9), e2023GL107647.

Li, L., Mahowald, N.M., Gonçalves Ageitos, M. et al. Improved constraints on hematite refractive index for estimating climatic effects of dust aerosols. Commun Earth Environ 5, 295 (2024). https://doi.org/10.1038/s43247-024-01441-4

Reply: *We thank the reviewer for the suggested references and have included them where appropriate.* **See line 85 -89**

Comment: Line 78-79: The term "pure dust" requires clarification. Do you mean particles composed solely of mineral components (though not limited to a single mineral species)? If so, this should be explicitly stated. Moreover, the claim that "pure dust is insoluble" is not strictly accurate. Some of the mineral components like hematite are indeed largely insoluble in water under ambient atmospheric conditions, but some others like calcite are highly soluble, and these are often present in atmospheric dust in small but non-negligible amounts.

Reply: *We agree that the term pure dust is not the most appropriate in the context of the study and have removed it. We agree that models assuming dust being "insoluble" is a simplification, ignoring the variable nature of dust particles. We have clarified in table 1 whether models have dust particles to act as CCN and further in the models description paragraphs which processes and assumptions led dust acting as CCN.*

Comment: Line 88-89: A reference is needed here, at minimum for K-feldspar. The following is one example that could be cited.

Atkinson, J. D., Murray, B. J., Woodhouse, M. T., Whale, T. F., Baustian, K. J., Carslaw, K. S., ... & Malkin, T. L. (2013). The importance of feldspar for ice nucleation by mineral dust in mixed-phase clouds. Nature, 498(7454), 355-358.

Reply: *In the revised manuscript this sentence has been removed.*

Line 103-105: The two sentences seem to contradict one another: "Furthermore, the net dust effective radiative forcing (DuERF) varies across models…" and "models may appear consistent in DuERF".
The first sentence states that DuERF varies between models, suggesting inconsistency, while the second says models "may appear consistent" which is the opposite.

Reply: *In the revised manuscript this sentence has been removed.*

Comment: Line 112: It would be helpful to specify which five models are being referred to.

Reply: *We appreciate the reviewers suggestion and have specified the models in the revised manuscript.* ***See Line 179-183***

Line 133-134: "wind dependent dust emission schemes". The current wording is somewhat informal, as dust emission in many ESMs depends on more than just wind speed. Additional factors such as the extent of bare soil, soil texture, and surface aridity also play critical roles in determining dust source regions.

Reply: *We agree with the reviewer's comment and have updated the text accordingly, see* ***Line 205- 206****.*

Line 150: Can the authors specify the composition of the mineral mixture used in this calculation?

Reply: *We thank the reviewer for the question. The internal mixture is between mineral dust and other aerosols: "When black carbon, dust, or both are present in the mix, these are treated as inclusions in a homogeneous background medium, using the Maxwell Garnett mixing rule." — von Noije et al. (2021). In other words, the statement does not refer to a mixture of minerals. EC-Earth3-AerChem uses a single set of refractive indices to represent mineral dust.*

Comment: Line 151: It would be useful to clarify how this model, along with others referenced, treats LW aerosol scattering.

Reply: *We thank the reviewer for the useful suggestions, we have included the treatment of LW scattering as a new column to Table 1.*

Comment: Line 191-192: In addition to the complex refractive index listed in Table 1, which specific optical properties are held fixed, e.g., single scattering albedo, extinction coefficient? Do these properties lack both spatial and temporal variability in the model?

Reply: *We thank the reviewer for pointing out ambiguities in the text. What we mean is that dust in the CNRM model is externally mixed and thus there is not any interaction between dust and other atmospheric species and accordingly the optical properties (SSA, MEC) of dust in each of the three bins do not change with time. In the revised manuscript we have clarified this.*

Comment: Line 269-271: This sentence should be revised for clarity or just removed. Since surface albedo is consistent across the models, it does not contribute to the spread in forcing efficiencies. This point is already clearly conveyed in the following sentence.

Reply: *We thank the reviewer for noticing this and we agree it was not a meaningful sentence, we have revised the sentence for clarity.* ***See line 360-366***

Comment: Line 313: Are there any insights into why these two models produce notably different results? Identifying specific parameterizations or assumptions that set them apart would strengthen the analysis.

Reply: *See reply to comment question 3. We now point to how UKESM1-0-LL is the only model representing LW scattering effects.*

Comment: Line 355: Why focus on AOD and AAOD rather than DOD and ADOD, which are more directly attributable to dust aerosols? Using dust-specific metrics would provide a clearer assessment of dust aerosol impacts, right?

Reply: *See our response to question* 6.

Comment: Line 362-363: The uncertainty range reported by Ridley et al. (2016) pertains to present-day DOD. It is unclear why the authors compare this with the modeled change in total AOD under preindustrial conditions.

Reply: *We appreciate the reviewer's question and hope this is partly clarified by our response to question 6. Since the dust emissions are the only factor changing, the difference in AOD between piClim-2xdust and piClim-control, representing a doubling of dust, should be a good approximation of the DOD in piClim-control. Furthermore the dust emissions in the models have not significantly changed between present day and pre-industrial. It is useful to compare the AOD perturbation (=preindustrial DOD) to that of the best available DOD of present day conditions. It also is useful to put in perspective the difference between piClim-2xdust and piCim-control comparing it to the present day dust effective radiative effect, especially for the direct effect, which changes rather linearly as a function of burden and optical depth.*

Comment: Line 379: I expect the authors to further propose, in addition to constraints on DOD, what other specific constraints should be included to reduce the uncertainty in the direct DuERF.

Reply: *We thank the reviewer for the suggestion and have added the following sentence* ***See line 505-508***

Going forward, we need to expose ESMs to a larger set of constraints on different aspects of the dust cycle, for example, particle size distribution (Kok et al., 2021), CRI (Li et al., 2024; Wang et al., 2024) or spatial gradients in DOD to constrain the lifetime of dust to reduce the uncertainty in the direct DuERF.

Comment: Line 480: Could the authors clarify the distinction between "dusty surface albedo" and the "planetary albedo effect from airborne dust"?

Reply: *We agree as written this was somewhat unambiguous, we have revised the text accordingly* ***See Line 645-649***

Comment: Line 517-518: Are there any references, figures, or tables provided to support this statement?

Reply: *We thank the reviewer for the question. We agree that the results we show in the manuscript do not offer sufficient evidence to support the claims of the statement as it was written, we have changed the wording of the statement and added a reference.* ***See Lines 697-698***