Reply to Reviewer #2

We would like to thank Reviewer #2 for the constructive and helpful comments. The reviewer's contribution is recognized in the acknowledgments of the revised manuscript. Below follows our response point by point. The reviewer's comments are given in *italic* and our response is given in **bold** font.

Major comments:

1) The Reviewer notes: "*1. The authors should more strongly emphasize novel aspects in their introduction, discussion, and conclusions and shorten the discussion of more general aspects. Some of the general discussions are rather lengthy while still missing essential points. These general discussions should be shortened.*"

The purpose and novelties of this paper are highlighted throughout the manuscript (e.g., Lines 175-184, Lines 218-220, Lines 265-267 and Lines 440-446). The novelty of this paper lies in the investigation of more technical aspects of ERF estimation, such as the robustness of the results, the relative contribution of ERFari, ERFaci and ERFalb to the total ERF, and the temporal evolution of ERFari, ERFaci and ERFalb in both global and regional scale (also see reply to Comment #5). This study reviews and complements the findings of previous studies by providing figures with ΔAOD and ERF spatial patterns and tables with weighted mean values and standard deviations on global and regional scale. Yet, indeed some parts of the introduction are lengthy, and are shortened in the revised manuscript as suggested.

2) The Reviewer notes: "*2. Thornhill et al. (2021) present a number of very similar results. I think that in order to justify another publication on this topic, the authors should try to come up with additional findings. Specifically, I suggest to analyze which regions contribute most strongly to the spread of model results for global mean ERFaci. I suggest to use a pragmatic separation into ΔAOD regimes based on multi-model average AOD change (ΔAOD) in Figures 1 a and b. For example, the authors could distinguish between high to medium ΔAOD source regions over land, medium to low ΔAOD regions over land, high ΔAOD over ocean in the outflow and low ΔAOD over ocean (remote ocean). I think it would be interesting not only to specify ERFaci standard deviations in W m^{-2} but also as contributions to uncertainty in global mean ERFaci, which involves area weighting.*"

Analysis of the ERF inter-model variability (one standard deviation) indicates that ERFaci is the main source of uncertainty in total ERF (Table I, Fig. I). East Asia contributes the most to the inter-model spread of both ΔAOD (Table II) and ERFaci results with the standard deviation of the latter exceeding 5.5 W m^{-2}. Tables I and II were incorporated inTables 5 and S5, respectively. Figure I was also added to the Supplement.

Table I. Inter-model variability (one standard deviation) of ERF (in W m^{-2}) during the negative ERF peak period (1965-1984) and the recent past (1995-2014) from the histSST experiment. Global and regional ERF standard deviations are presented for the five regions investigated in the paper: East North America (ENA), West and Central Europe (WCE), the Mediterranean (MED), East Asia (EAS) and South Asia (SAS).

Region	1965-1984				1995-2014			
	ERF	ARI	ACI	ALB	ERF	ARI	ACI	ALB
ENA	1.61	0.59	1.84	0.15	1.00	0.39	1.23	0.15
WCE	3.00	0.59	3.65	0.51	1.13	0.24	1.35	0.49
MED	0.96	0.45	1.40	0.30	0.61	0.31	1.18	0.52
EAS	2.45	0.46	2.90	0.25	3.06	1.03	3.71	0.18
SAS	0.89	0.31	0.97	0.46	0.91	0.86	1.76	0.92
GLOBAL	0.43	0.11	0.44	0.05	0.37	0.14	0.44	0.08

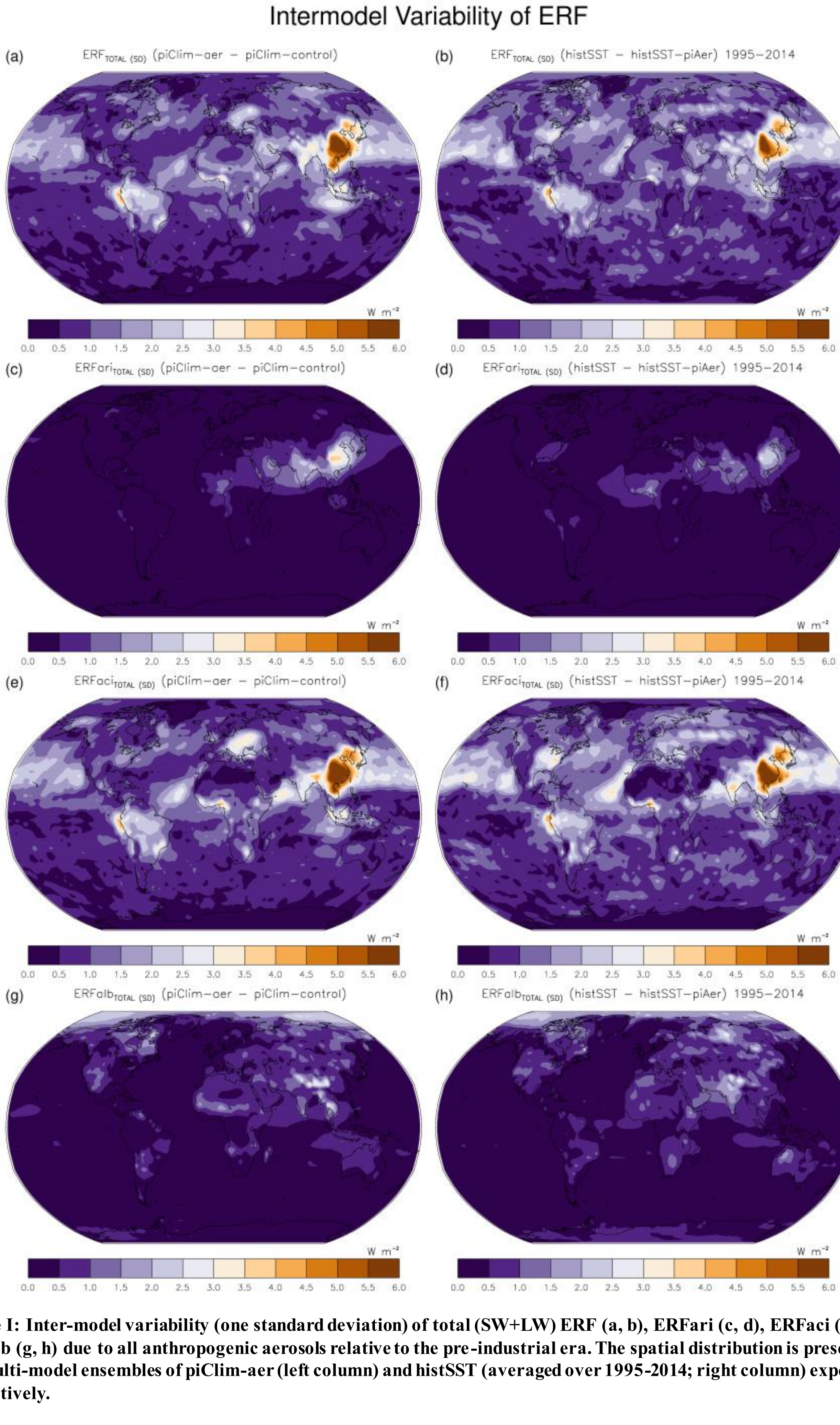

Figure I: Inter-model variability (one standard deviation) of total (SW+LW) ERF (a, b), ERFari (c, d), ERFaci (e, f) and ERFalb (g, h) due to all anthropogenic aerosols relative to the pre-industrial era. The spatial distribution is presented for the multi-model ensembles of piClim-aer (left column) and histSST (averaged over 1995-2014; right column) experiments, respectively.

Table II. Inter-model variability (one standard deviation) of ΔAOD for histSST experiment averaged over 1965-1984 and 1995-2014. Variables od550aer, od550so4, od550oa, and od550bc denote the differences in all-aerosol, sulphate, organic aerosol, and black carbon AOD, respectively. Global and regional ΔAOD standard deviations are presented for the five regions investigated in the paper: East North America (ENA), West and Central Europe (WCE), the Mediterranean (MED), East Asia (EAS) and South Asia (SAS).

Region	1965-1984				1995-2014			
	od550aer	od550so4	od550oa	od550bc	od550aer	od550so4	od550oa	od550bc
ENA	0.0418	0.0224	0.0095	0.0006	0.0204	0.0168	0.0080	0.0006
WCE	0.1182	0.0878	0.0076	0.0015	0.0310	0.0276	0.0046	0.0008
MED	0.0258	0.0191	0.0035	0.0005	0.0162	0.0129	0.0033	0.0006
EAS	0.0340	0.0262	0.0086	0.0017	0.0839	0.0649	0.0216	0.0039
SAS	0.0127	0.0085	0.0069	0.0008	0.0495	0.0309	0.0187	0.0015
GLOBAL	0.0065	0.0068	0.0023	0.0003	0.0088	0.0077	0.0039	0.0005

Based on Fig. 1a and 1b of the manuscript, four ΔAOD regimes can be distinguished:

a) **High to medium ΔAOD over land: East and South Asia, Eastern Europe, Middle East, Eastern North America**
b) **Medium to low ΔAOD over land: North and South America, Western Europe, Greenland, Oceania, Antarctica, Arctic**
c) **High ΔAOD over ocean: Northwestern Pacific, Northernmost Indian**
d) **Low ΔAOD over ocean: Atlantic, South Pacific, South Indian**

The above ΔAOD regime discussion was added to the "Results" section.

3) The Reviewer notes: "*3. It is widely understood that emissions took different trajectories in Europe and Asia and also that the trajectories for India and China have diverged. I think the authors should either shorten or omit the analysis of selected regions or else explain better what motivated this part of the study and what is new or unexpected about their results.*"

The time evolution of ΔAOD and all ERF components (ARI, ACI and ALB) is shown on global scale and over ERF hotspots during 1850-2014 using histSST experiments. The selected regions are highly industrialized regions that exert the most negative ERF values during the historical period. Weighted field means of ΔAOD and ERF over the selected regions are also presented for every ESM and their ensemble for two time periods of interest: i) during the negative ERF peak (1965-1984; see also Szopa et al., 2021) and ii) the end of historical simulations (1995-2014). Moreover, SW and LW ERF values are presented over the selected regions for both time periods of interest, which not only addresses the errata of IPCC AR6 Chapter 6 (Szopa et al., 2021; https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_Errata.pdf), but also attempts to explain the underlying physical processes by studying changes in liquid and ice water paths (Fig. S10 in Supplement).

4) The Reviewer notes: "*4. Lines 537-556: I think that in the conclusion section the authors should summarize results and draw conclusions instead of spending an entire long paragraph simply repeating what they did.*"

The reviewer is right. As a result, the "Conclusions" section was modified (see reply to Comment #5).

5) The Reviewer notes: "*5. Lines 536-611: Novel findings should be highlighted. If the only main points are to demonstrate that the authors used similar methods to arrive at similar results compared to previous studies, then I do not understand why this study presents an advance over previous studies and should be published.*"

The authors of this paper strongly believe that this study presents a number of novel findings as it is a comprehensive spatiotemporal ERF analysis, which complements and advances the findings of other papers:

a) **A concise ensemble of seven CMIP6 Earth System models was used to calculate the ERF of anthropogenic aerosols (as a whole and for different sub-species, such as black and organic carbon, and sulphates) from two different sets of experiments (piClim series and histSST) based on the method of Ghan (2013), which is considered a very accurate method (Zelinka et al., 2014; Michou et al., 2020).**
b) **Spatial patterns at top-of-atmosphere and global weighted field means for all SW, LW and total (i.e., SW+LW) ERF components (ARI, ACI and ALB) are presented for every ESM and their ensemble and for every experiment (piClim-aer, piClim-BC, piClim-OC, piClim-SO_2 and histSST). Inter-model variability (one standard deviation) is also shown in the case of the ensemble. To our knowledge, the information obtained from a) and b) has not been presented all together in one paper.**
c) **The inter-model agreement as well as the robustness of ERF and ΔAOD ensemble results were calculated using a method similar to the one used in the 6th IPCC Assessment Report (described within the manuscript) and are shown in Fig. 1-3. This method of calculating the robustness of the results has not been used in other papers as far as we know.**
d) **The novel concept of determining the driving factor of ERF (ARI, ACI or ALB) on global scale is presented in Fig. 6-7 using a method described in detail in the text. This has not been done in other studies.**

Although the novelty of this study is highlighted in the last paragraph of the introduction, the reviewer has a point. Considering Comments #4 and #5, the first paragraph in the "Conclusions" section was modified as follows: "In this work, the effective radiative forcing (ERF) of anthropogenic aerosols was investigated using fixed-SST simulations from seven different ESMs participating in the CMIP6 exercise. Shortwave (SW), longwave (LW) and total (i.e., SW+LW) ERF and changes in aerosol optical depth (AOD) were quantified for all anthropogenic aerosols, combined and individually, using both piClim and histSST experiments for comparison purposes. Additionally, the robustness of the multi-model ensemble results was calculated by investigating both the statistical significance of each model's results and the agreement between individual models on the sign of change. Spatial patterns and temporal evolution of ERF and ΔAOD were presented on global and regional scale, along with tables that show the area-weighted mean values and standard deviation of ERF and ΔAOD for the multi-model ensemble as well as every individual model."

Specific comments:

6) The Reviewer notes: "*Line 23: I think that "which is the recommended metric for perturbations affecting the Earth's top-of-atmosphere energy budget since it is a better way to link this perturbation to subsequent global mean surface temperature change" and also the corresponding lengthy discussion in lines 106 to 130 of the introduction should be omitted. The question whether ERF is a good metric is not addressed by the results of this study.*"

The reviewer is correct. Lines 23-24 and Lines 110-130 were omitted.

7) The Reviewer notes: "*Line 62: "can be observed" -> please elaborate*"

This sentence merely addresses the fact that a semi-direct effect of aerosols exists and has been investigated in a number of papers mentioned in the manuscript. Although it is not investigated individually, it is part of the ERFaci term of Ghan's (2013) decomposition and for reasons of completeness, it is discussed in the introduction.

8) The Reviewer notes: "*Line 214: I think that for an ERF it is sufficient that identical SST and SIC are prescribed in the base and the perturbed run. The (first order) question is whether SST and SIC are allowed to respond to the forcing.*"

The histSST and histSST-piAer simulations use atmosphere-only configurations with prescribed sea-surface temperatures and sea ice (Collins et al., 2017). Therefore, SSTs and SIC are not allowed to respond to the aerosol forcing.

9) The Reviewer notes: "*Lines 218-220: Perhaps explain and motivate this in the introduction?*"

Explaining this in the introduction would be confusing for the readers, as the simulations used in this paper are described in Chapter 2. However, the following sentences were added to the last paragraph of the introduction:

"The present-day anthropogenic aerosol ERF is examined at the top-of-the-atmosphere using two different sets of experiments with fixed sea surface temperatures (SSTs) and sea ice cover (SIC) for comparison purposes. Moreover, the evolution of transient ERF during the historical period (1850-2014) is investigated globally and over certain emission regions of the Northern Hemisphere (NH), focusing on the last 20 years of the historical period (1995-2014) in order to mitigate the effects of the negative ERF peak around in late 1970s (Szopa et al., 2021)."

10) The Reviewer notes: "*Lines 378-381: What do we learn from this? If I interpret it correctly, Figure 4 suggests that the values are consistent.*"

Lines 378-381 show the similarities between piClim-aer and histSST (averaged over 1995-2014), but also highlight the differences between ESMs in ERFari and ERFaci. While all models agree on the negative sign of ERFaci for both experiments, there are discrepancies in the sign of ERFari. These statements were also added in the revised manuscript.

11) The Reviewer notes: "*Lines 454-455: Why?*"

Due to differences in experimental set-up between piClim-aer and piClim-SO_2, as well as differences in the number of models used for the ERF decomposition (EC-Earth3-AerChem did not participate in the calculation of ERF for piClim-SO_2, piClim-OC and piClim-BC due lack of diagnostics as stated in Lines 262-263). One possible reason is that some of the cooling by sulphate aerosols is compensated by warming by BC in the piClim-aer experiment compared to piClim-SO_2. However, it should also be noted that, while the method used to determine the dominant ERF component can provide some insights, it is quite simplistic and, therefore, cannot fully explain the underlying physical processes that lead to the forcing.

12) The Reviewer notes: "*Table 4: For MPI-ESM-1-2-HAM piClim, the sum of ERFs for individual species seems to differ more from the total ERF than for other models. Do you have idea why this could be?*"

MPI-ESM-1-2-HAM produces a highly negative ERFaci in both piClim-OC and piClim-BC, thus leading to more negative total ERF for both experiments. As a result, the sum of ERFs for individual aerosol species differs more from the ERF due to all anthropogenic aerosols. The reason behind these results is that coating by anthropogenic sulphate removes BC and OC in piClim-aer, whereas OC and BC have longer lifetimes in piClim-OC and piClim-BC respectively (not shown). Therefore, OC and BC are transported further and contribute more to ERF in piClim-OC and piClim-BC respectively than in piClim-aer.

Technical comments:

13) The Reviewer notes: "*Line 399: borne to mind -> borne in mind?*"

It was revised accordingly as suggested by the reviewer.

14) The Reviewer notes: "*Lines 399-400: "affecting the realizations and parameterization schemes ESMs use to quantify the magnitude of different processes" sound confusing. Please omit or rephrase. Avoid unnecessary repetitions.*"

Lines 398-401 were rephrased as follows: "It should be borne in mind that not all ESMs agree on the magnitude or even the sign of the individual SW and LW ERF main components (Tables S2-S4) due to uncertainties in the parameterization schemes used in ESMs to describe the way aerosols interact with radiation and clouds."

15) The Reviewer notes: "*Line 405: It is interesting to note that -> The*"

It was revised accordingly as suggested by the reviewer.

16) The Reviewer notes: "*Line 454: larger -> a larger*"

It was revised accordingly as suggested by the reviewer.

17) The Reviewer notes: "*Line 498: gets more positive values -> becomes less negative*"

It was revised accordingly as suggested by the reviewer.

References

Collins, W. J., Lamarque, J.-F., Schulz, M., Boucher, O., Eyring, V., Hegglin, M. I., Maycock, A., Myhre, G., Prather, M., Shindell, D., and Smith, S. J.: AerChemMIP: quantifying the effects of chemistry and aerosols in CMIP6, Geosci. Model Dev., 10, 585–607, https://doi.org/10.5194/gmd-10-585-2017, 2017.

Michou, M., Nabat, P., Saint-Martin, D., Bock, J., Decharme, B., Mallet, M., Roehrig, R., Séférian, R., Sénési, S., and Voldoire, A.: Present-Day and Historical Aerosol and Ozone Characteristics in CNRM CMIP6 Simulations, J. Adv. Model. Earth Syst., 12, https://doi.org/10.1029/2019MS001816, 2020.

Szopa, S., Naik, V., Adhikary, B., Artaxo Netto, P. E., Berntsen, T., Collins, W. D., Fuzzi, S., Gallardo, L., Kiendler-Scharr, A., Klimont, Z., Liao, H., Unger, N., and Zanis, P.: Short-lived climate forcers, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, Ö., Yu, R., and Zhou, B., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 817–922, https://doi.org/10.1017/9781009157896.001, 2021.

Zelinka, M. D., Andrews, T., Forster, P. M., and Taylor, K. E.: Quantifying components of aerosol-cloud-radiation interactions in climate models, J. Geophys. Res. Atmos., 119, 7599–7615, https://doi.org/10.1002/2014JD021710, 2014.