Response to the comments of reviewer #3

I commend the authors on a much improved manuscript. The significant investment of effort into a sentence-by-sentence rewriting of the manuscript by the entire author list has made the conclusions much easier for the reader to grasp. As a minor comment, since "larger audience" is mentioned multiple times in the authors' reply, I do want to clarify that the issue was never that the reader had to be an expert, but rather that even the experts were left confused.

My very minor comments use the differences document attached to the authors' responses because this was the only document I had access to at the time due to computer problems while traveling.

We thank the reviewer for the appraisal. We are glad that our effort in revising the language helped making the text clearer. In the following, we reply (blue) to your comments (black).

p. 4, l. 10: I am still unclear on the definition of Fari (and aci). From the author replies, I am led to believe these are not radiative flux differences, but rather . . . what? The physical processes themselves?

We changed it to: „(…) to induce radiative effects associated with the physical processes of aerosol-radiation interactions (F_{ari}) and aerosol-cloud interactions (F_{aci}) in a consistent manner."

p. 5, l. 25: what does "interpretation" mean?

Changed to: „MACv2 is also based on more recent emission data relative to 1850 (…)"

p. 6, l. 18: I know what you mean, but you might want to rewrite this sentence to make it easier to understand why this is efficient;

We add the reason:„ The six pre-industrial simulations serve as the reference for the experiments with anthropogenic aerosol and therefore efficiently increase the number of forcing estimates for anthropogenic aerosol."

on the previous line, "three experiment" ! "three experiments"

Changed as suggested.

p. 7, l. 13: state explicitly that "ADJ" is the semidirect effect except for EC-EARTH.

Added: „Our rapid adjustments are associated with atmospheric temperature changes, i.e., semi-direct effects, except for EC-Earth accounting also for adjustments in cloud microphysics."

I am a little confused that the adjustments are positive; I thought that they were typically negative in GCMs. Is this because the Twomey effect is weaker than the direct aerosol effect in MAC–SP (which I think is also unusual)? It might be good to provide a few sentences of context for these MAC–SP results compared to other multi-model ensembles (AeroCom, CMIP5); I think most readers will be convinced that they should expect lower intermodel spread when the Twomey effect is prescribed, but they may share my surprise when effects change sign compared to the conventional wisdom, i.e., all-sky effect stronger than clear-sky, semidirect effect negative).

When MACv2-SP is implemented in the radiation calculation, the net contribution from rapid adjustment is positive. This is the case for all our models except EC-Earth and consistent with buffering of perturbations in the atmosphere. Most GCMs, however, perturb also the cloud microphysics. This is expected for models implementing MACv2-SP in the cloud microphysics like

EC-Earth that induces also rapid adjustments in cloud processes. Such negative adjustments are consistent with your perception of GCM results. These can be large in models, but their magnitudes remain uncertain (Bellouin et al., in prep.).

The Twomey effect from the standard setup of MACv2-SP in ECHAM is about the same size as the direct radiative forcing, namely -0.27Wm^{-2} and -0.23Wm^{-2} (Fiedler et al., 2017). This relative contribution is in agreement with the last assessment report by the IPCC (Myhre et al., 2013). It has been previously discussed for MACv2-SP in Fiedler et al. (2017) alongside an assessment of the impact of the parameterization of the Twomey effect and a discussion of a more negative present-day forcing of -1Wm^{-2} when a stronger Twomey effect is induced with MACv2-SP. The magnitude of the Twomey effect is therefore easy to adapt with MACv2-SP, when the magnitude will be better constrained, but it currently remains uncertain (Bellouin et al., in prep.).

We added in Section 3.3: „The positive net contribution from adjustments is consistent with buffering of perturbations by atmospheric processes. A discussion of the rapid adjustments and the choice for the Twomey effect in ECHAM is given by Fiedler et al. (2017).“ The references to previous inter-comparisons from Myhre et al. (2013) and Shindell et al. (2013) are added in Section 3.1: „One could expect less model diversity in all-sky ERF from our study than from previous inter-comparison projects (e.g., Myhre et al., 2013, Shindell et al., 2013), because we prescribe the same aerosol optical properties and the associated change in cloud droplet numbers.“

p. 7, l. 21: remove "possible"
Removed.

p. 8, l. 3: it might be good to add a little more discussion of these results in light of the prevailing opinion that preindustrial/background aerosol properties constitute a large (and irreducible) uncertainty on the anthropogenic forcing; perhaps this result suggests that this problem is not so severe?
We added: „Using different parameterizations for Faci can change this result because of non-linear processes. The magnitude of Faci, however, remains uncertain (Bellouin et al., in prep.). One contributing uncertainty is the poor quantitative understanding of the pre-industrial aerosols (e.g., Carslaw et al., 2013).“

p. 8, l. 7: "to" ! "from"
Changed as suggested.

p. 8, l. 12: In light of this statement, it would be interesting to see the clear-sky values added to Tab. 2. If the Twomey and direct effects are relatively close in MAC–SP, then it would not be surprising that cloud parameterization doesn't affect the forcing.
Table 2 lists the clear-sky ERFs in brackets, illustrating the model diversity in both all- and clear-sky ERF from MACv2-SP. We added: „One could expect less model diversity in all-sky ERF from our study than from previous inter-comparison projects (e.g., Myhre et al., 2013, Shindell et al., 2013), because we prescribe the same aerosol optical properties and the associated change in cloud droplet numbers. However, our model diversity in clear-sky ERF is smaller than for our all-sky ERF (Table 2). This points to the influence of model differences in representing clouds (Appendix B) on the all-sky ERF. Our results therefore indicate that model differences in

meteorological parameters contribute to the model diversity in all-sky ERF. This is also the case for the ERF uncertainty in a complex aerosol-climate model (Regayre et al., 2018)." Please refer to our response aloft for the relative contributions from the Twomey and direct effects in ECHAM.

Sec. 3.5: I like this additional text very much. I believe you should rephrase the sentence starting on p. 12, l. 7 ("A clear saturation. . . ") to reflect more clearly that the largest efficiencies occur at the (arbitrarily chosen) edges of the plumes, one of the reasons we would expect MAC–SP to provide a lower bound on the strength of the Twomey effect. (See my first comment on the previous manuscript version.)

We had added the new figure 9 for illustrating that the largest efficiencies do not necessarily occur at the edges of the plumes. The text is now: „Large efficiencies occur typically in remote areas including some regions at the edges of τ_a plumes (Fig. 9). No clear saturation of Faci is evident at all edges of the τ_a plumes. Also the spatial distribution of both the all- and cloudy-sky efficiency is rather inhomogeneous."

The position and extent of the plumes (Stevens et al., 2017) were chosen for reproducing the spatial distributions of the monthly mean in the anthropogenic aerosol optical depth from the MPI-M Aerosol Climatology that has changed between version 1 and version 2 (MAC, Kinne et al., 2013, Kinne, in review, Section 2.1 in the manuscript). Please also note our comment aloft on the uncertainty of the magnitude of the Twomey effect (Bellouin et al., in prep.) as well as the flexibility to induce different strengths and patterns for the Twomey effect with MACv2-SP (Fiedler et al., 2017).

p. 13, l. 19: I thank the authors for humoring my comment on nudging (which I realize in retrospect made no sense in the context of their methods). I think it is good to include this conclusion, and I agree with it except for the part about interfering with adjustments. For example, Ghan et al. (2016), of which some of this study's authors are also coauthors, derive adjustments based on nudged simulations.

There seems to be a debate on the usage of nudging for diagnosing aerosol forcing in the community. Also Ghan et al. (2016) acknowledges that nudging temperatures influences the development of moist convection in some models based on the study by Zhang et al. (2014) and discourages the usage of this type of nudging. We added here: „The interference of nudging with adjustments deserves closer attention in future research."

Figure 4 b: update the label to "interannual"

Added „interannual" in both the label and the caption: „(…) as measure for the interannual variability internal to the model ensemble."

Anthropogenic aerosol forcing - insights from multiple estimates from aerosol-climate models with reduced complexity

Stephanie Fiedler[1], Stefan Kinne[1], Wan Ting Katty Huang[2], Petri Räisänen[3], Declan O'Donnell[3], Nicolas Bellouin[4], Philip Stier[5], Joonas Merikanto[3], Twan van Noije[6], Risto Makkonen[3,7], and Ulrike Lohmann[2]

[1]Max Planck Institute for Meteorology, Hamburg, Germany
[2]Institute for Atmospheric and Climate Science, ETH Zürich, Zürich, Switzerland
[3]Finnish Meteorological Institute, Helsinki, Finland
[4]Department of Meteorology, University of Reading, Reading, UK
[5]Department of Physics, University of Oxford, Oxford, UK
[6]Royal Netherlands Meteorological Institute, De Bilt, Netherlands
[7]Institute for Atmospheric and Earth System Research / Physics, Faculty of Science, University of Helsinki, Finland

Correspondence to: Stephanie Fiedler (stephanie.fiedler@mpimet.mpg.de)

Abstract. This study assesses the change in anthropogenic aerosol forcing from the mid-1970s to the mid-2000s. Both decades had similar global mean anthropogenic aerosol optical depths, but substantially different global distributions. For both years, we quantify (i) the forcing spread due to model internal variability and (ii) the forcing spread among models. Our assessment is based on new ensembles of atmosphere-only simulations with five state-of-the-art Earth system models. Four of these models

5　will be used in the sixth coupled model inter-comparison project (CMIP6, Eyring et al., 2016). Here, the complexity of the anthropogenic aerosol has been reduced in the participating models. In all our simulations, we prescribe the same patterns of the anthropogenic aerosol optical properties and associated effects on the cloud droplet number concentration. We calculate the instantaneous radiative forcing (RF) and the effective radiative forcing (ERF). Their difference defines the net contribution from rapid adjustments. Our simulations show a model spread in ERF from -0.4 to -0.9 Wm^{-2}. The standard deviation in annual

10　ERF is 0.3 Wm^{-2}, based on 180 individual estimates from each participating model. This result implies that identifying the model spread in ERF due to systematic differences requires averaging over a sufficiently large number of years. Moreover, we find almost identical ERFs for the mid-1970s and mid-2000s for individual models, although there are major model differences in natural aerosols and clouds. The model-ensemble mean ERF is -0.54 Wm^{-2} for the pre-industrial era to mid-1970s and -0.59 Wm^{-2} for the pre-industrial era to mid-2000s. Our result suggests that comparing ERF changes between two observable

15　periods rather than absolute magnitudes relative to a poorly constrained pre-industrial state might provide a better test for a model's ability to represent transient climate changes.

1　Introduction

Despite decades of research on the radiative forcing of anthropogenic aerosol, quantifying the present-day magnitude and reconstructing the historical change of the forcing remains challenging. Figure 1 shows the anthropogenic aerosol optical

depth for the mid-1970s and mid-2000s that we use in this study (Fiedler et al., 2017; Stevens et al., 2017). The anthropogenic aerosol pollution in the mid-1970s was larger in Europe and North America than in East Asia, whereas the opposite is the case in the mid-2000s. In addition to these regional changes in aerosol pollution, differences in the surface albedo, insolation, and cloud regimes between the aerosol transport regions of the Pacific and continental Europe may result in temporal changes in the global effective radiative forcing (ERF). Based on a single state-of-the-art climate model, the long-term and global ERF does not change despite the substantial spatial changes in anthropogenic aerosol optical depth (τ_a) between the mid-1970s and mid-2000s (Fiedler et al., 2017). Internal model variability, however, strongly affects annual estimates of the global mean effective radiative forcing.

In light of model uncertainties (e.g., Kinne et al., 2006; Quaas et al., 2009; Lohmann and Ferrachat, 2010; Lacagnina et al., 2015; Koffi et al., 2016), the use of a single model does not necessarily represent the full spectrum of possible anthropogenic aerosol forcings. In the present study, we therefore revisit the question of Fiedler et al. (2017): "Does the substantial spatial change of the anthropogenic aerosol between the mid-1970s and mid-2000s affect the global magnitude of ERF?", based on ensembles of simulations from five global aerosol-climate models, all using identical anthropogenic aerosol perturbations of reduced complexity. In this context, we additionally ask: "What is the relative contribution of internal model variability to the ERF spread?", and document the model diversity for the pre-industrial aerosol as well as cloud characteristics and the surface albedo that are relevant for the ERF of anthropogenic aerosol. Such model differences have previously been identified for other climate models (e.g., Stier et al., 2007; Nam et al., 2012; Fiedler et al., 2016; Crueger et al., 2018).

Previously a reduction of the model complexity has been accomplished by prescribing idealized aerosol radiative properties, e.g., within the framework of Aerosol Comparisons between Observations and Models (AeroCom, e.g., Randles et al., 2013; Stier et al., 2013). Here, we prescribe observationally constrained optical properties of anthropogenic aerosol and an associated effect on the cloud droplet number concentration with the simple plumes parameterization (MACv2-SP, Fiedler et al., 2017; Stevens et al., 2017), but keep the full model diversity in all other aspects. The approach eliminates uncertainties in process modelling of anthropogenic aerosol such that our study represents uncertainties associated with other processes influencing the radiative forcing. In other words, by using MACv2-SP in the participating models, the model inter-comparison allows us to investigate those sources of uncertainty that remain if we pretend to know the spatial distribution of anthropogenic aerosol. This work can be seen as a pilot study for the "Radiative Forcing Model Inter-comparison Project" (RFMIP, Pincus et al., 2016), endorsed by CMIP6 (Eyring et al., 2016), using the same experiment setup with MACv2-SP.

Throughout our model inter-comparison, we consider the effect of model-internal variability on estimates of ERF. We do so by producing equally-sized ensembles of simulations for all participating models. Model-internal variability in this context is defined as the year-to-year changes in model parameters associated with inter-annual variations of the meteorological state. The results of the climate models are compared with satellite data and a stand-alone radiative transfer model. The following section introduces the models and the experiment strategy in more detail, followed by our discussion of the results in Section 3 and conclusions in Section 4.

2 Method

2.1 Participating Models

This work uses five Earth system models and one stand-alone radiative transfer code. The participating climate models, which are run here in an atmosphere-only mode, are the atmosphere component ECHAM6.3 of the Earth system model MPI-ESM1.2 (Mauritsen et al., 2019) of the Max-Planck Institute for Meteorology (MPI-M), ECHAM6.3-HAM2.3 from the ETH Zürich (Tegen et al., 2018; Neubauer et al., submitted to ACPD), EC-Earth (e.g., Hazeleger et al., 2010; Döscher et al., in prep.) run at the Royal Netherlands Meteorological Institute, NorESM (Bentsen et al., 2013; Iversen et al., 2013; Kirkevåg et al., 2013) run at the Finnish Meteorological Institute, and HadGEM3 (Walters et al., 2017) developed at the UK Met Office. All models except ECHAM6.3 can treat aerosols and their interaction with meteorological processes with complex process-based parameterisation schemes linking aerosols to radiation and clouds. ~~All~~ In this study, all physics packages except the parameterization of anthropogenic aerosols are model-dependent, e.g., the treatment of the pre-industrial aerosols and clouds differ. Appendix A summarizes differences in radiation, cloud, and aerosol physics packages of the participating models.

In the present study, we prescribe the distributions of anthropogenic aerosols in all models following the MACv2-SP approach (Fiedler et al., 2017; Stevens et al., 2017). MACv2-SP mimics the spatio-temporal distribution and wavelength dependence of the optical properties of anthropogenic aerosols as well as a change in the cloud droplet number concentration (N) to ~~represent~~ induce radiative effects associated with the physical processes of aerosol-radiation interactions (F_{ari}) and aerosol-cloud interactions (F_{aci}) in a consistent manner. To do so, MACv2-SP uses analytical functions for approximating the monthly distribution of the present-day anthropogenic aerosol optical depth and the vertical profile of the aerosol extinction from the updated MPI-M aerosol climatology (MACv2, Kinne et al., 2013, Kinne et al, in prep.). Figure 1 shows the annual mean patterns of the anthropogenic aerosol optical depth (τ_a), and the fractional increase in the cloud droplet number concentration (η_N) relative to the pre-industrial level of 1850 from MACv2-SP. By design, MACv2-SP does not simulate sub-monthly variability in anthropogenic aerosol. Absorption of anthropogenic aerosol is prescribed with a mid-visible single scattering albedo of 0.93 for industrial plumes and 0.87 for plumes with seasonally active biomass burning. The anthropogenic aerosols are assumed to be small in size with an Angstrom parameter of 2 and an asymmetry parameter of 0.63. Here, we use MACv2-SP with the CMIP6 reconstructed changes of anthropogenic aerosol emissions, identical to the one used by Fiedler et al. (2017). Stevens et al. (2017) describe the technical details of MACv2-SP.

The use of the optical properties from MACv2-SP yields a consistent description of F_{ari}, including both direct radiative and semi-direct effects, across the models. All models account for the first indirect or Twomey effect by multiplying their cloud droplet number concentrations, calculated for pre-industrial aerosol conditions, by η_N prior to the radiative transfer calculation. Since η_N is larger than one in the presence of anthropogenic aerosols, the effective radius of cloud droplets is reduced, which enhances the cloud reflectivity of shortwave radiation. Note that η_N is only available for regions with $\tau_a > 0$ (see Fig. 1). In addition, the EC-Earth model also includes a second indirect or cloud lifetime effect by using the modified cloud droplet number concentrations in the cloud microphysics scheme (Döscher et al., in prep.).

We do not prescribe the same natural aerosol nor interfere with any other model components than prescribing the optical properties of anthropogenic aerosols and η_N. For instance, the pre-industrial aerosol optical depth (τ_p) depends on the model (Figures 2 and 3), which only affects F_{ari} and not F_{aci} as the prescription of η_N is identical in the participating models. Regional differences in τ_p occur primarily over oceans and deserts, where observations are typically sparse. It is noteworthy that ECHAM-HAM runs with interactive parameterisations for dust and sea-salt aerosol resulting in different spatio-temporal variability in τ_p (Figure 3), while in ECHAM the monthly climatology from MACv1 is prescribed. In the interactive parameterisations, the natural aerosol emissions, transport and deposition rely on meteorological processes that are difficult to represent in coarse-resolution climate models, e.g., desert-dust emissions strongly depend on the model representation of near-surface winds (e.g., Fiedler et al., 2016) such that constraining the desert-dust burden remains challenging in aerosol modelling (e.g., Räisänen et al., 2013; Evan et al., 2014; Huneeus et al., 2016). The aerosol-climate models also contain some anthropogenic aerosol in τ_p, but the majority of the pre-industrial aerosol optical depth is of natural origin. For instance, the 1850's global-mean τ_p in NorESM is 0.096, to which anthropogenic fossil-fuel aerosols contribute 0.002. For comparison, the here prescribed global mean τ_a is 0.029 for 2005.

In addition to the complex climate models listed above, we use the offline radiative-transfer model of Kinne et al. (2013) for an assessment of the instantaneous radiative forcing. This model has eight solar and twelve infrared bands, and reads monthly maps of the atmospheric and surface properties. These are for instance monthly means for the cloud properties from ISCCP and the surface albedo from the satellite product MODIS-SSM/I (Kinne et al., 2013). The radiative transfer calculation considers nine different sun elevations and eight randomly chosen combinations of cloud heights and overlap. The aerosol column properties at 550 nm are defined by the MPI-M Aerosol Climatology (MAC). The aerosol vertical distribution and the fine-mode anthropogenic fraction of aerosol optical depth for the mid-2000s are derived from global models participating in AeroCom (e.g., Myhre et al., 2013). We calculate the radiation transfer with both MAC version one (MACv1, Kinne et al., 2013) and two (MACv2, Kinne, in review). The latter considers more recent observational data, e.g., from the Maritime Aerosol Network (MAN, Smirnov et al., 2009), and a smaller anthropogenic aerosol fraction. MACv2 is also based on ~~a more recent AeroCom interpretation of the present-day~~ more recent emission data relative to 1850 (Lamarque et al., 2010), while MACv1 used emission data relative to 1750 (Dentener et al., 2006). The two climatologies therefore make different assumptions on the pre-industrial background, shown in Figure 3. The temporal scaling of anthropogenic aerosol optical depth in MACv1 and MACv2 is from the same transient ECHAM simulation (Stier et al., 2006). The parameterization form of the Twomey effect for MACv1 and MACv2 are here identical to MACv2-SP, but the assumptions for τ_p and τ_a differ.

2.2 Experiment strategy

All climate model simulations are carried out with the atmosphere-only configurations using prescribed monthly mean sea-surface temperatures and sea ice. Table 1 summarises the major characteristics of the model simulations. The modelling groups were free to set up all model components other than MACv2-SP, and choose their own boundary and initialization data. Specifically, the modelling groups use their own representation of pre-industrial aerosol for 1850 such that the present work includes

both models with prescribed monthly climatologies and interactive parameterisation schemes for natural aerosol species (Appendix A). Moreover, the physical parameterisations of radiation and clouds are different across the models (Appendix A).

Motivated by the effect of natural variability on ERF estimates in ECHAM (Fiedler et al., 2017), each model was run to produce a number of simulation ensembles: a reference ensemble consisting of six simulations with only pre-industrial aerosols representative for 1850, and two additional ensembles consisting of three simulations each with aerosols representative for 1975 and 2005, respectively. For each model, we perform a total of twelve experiments for the years $2000 - 2010$ inclusive. These are six experiments with τ_p for the year 1850, three experiments with τ_p and anthropogenic aerosol from MACv2-SP for the year 1975, and three ~~experiment~~ experiments with τ_p and anthropogenic aerosol from MACv2-SP for the year 2005. ~~Six~~ The six pre-industrial simulations ~~are chosen for efficiently increasing~~ serve as the reference for the experiments with anthropogenic aerosol and therefore efficiently increase the number of forcing estimates for anthropogenic aerosol. The first year of each run is considered as a spin-up period and is excluded from the analysis. A ten-year period was chosen to account for variability in the boundary conditions.

The instantaneous radiative forcing (RF) of anthropogenic aerosols in clear and all sky is estimated from double radiation calls in the models having this functionality, namely ECHAM, ECHAM-HAM and NorESM. Aerosol radiative effects predominantly occur for shortwave radiation. We therefore calculate the atmospheric transfer of shortwave radiation once with and once without the contribution from anthropogenic aerosols to the aerosol optical properties and their effect on the cloud droplet number concentration. For each model, this gives us in total 30 annual estimates of RF for each of the two τ_a patterns shown in Figure 1, which is sufficient to estimate the mean RF and can be directly compared to the offline radiation-transfer calculations. We calculate RF at the top of the atmosphere (TOA) and at the surface (SFC) and list the global means in Table 2.

The ERF is calculated as the difference in the shortwave radiative flux at the top of the atmosphere between the simulations with and without anthropogenic aerosols. For illustrating the effect of year-to-year variability, we calculate annual ERF estimates for each of the ten simulation years. Combining the six pre-industrial experiments with each of the three experiments with additional anthropogenic aerosol thus yields 6x3 annual ERF estimates for each year of the simulation, i.e., 180 annual estimates per model and τ_a pattern in total. We calculate the standard deviation from these 180 annual ERF values and use it as a measure of the natural variability in ERF internal to the models. The means of these 180 values are used for identifying systematic model differences in ERF. It was shown in an earlier study using ECHAM (Fiedler et al., 2017) that the combination of ensemble size and simulation length adopted here is sufficient for precisely estimating the ERF of a model. For comparison, the RFMIP protocol recommends a thirty-year average for diagnosing the ERF of a model (Pincus et al., 2016). Finally, we calculate the net contribution of rapid adjustments (ADJ) to ERF by subtracting RF from ERF for each model. Our rapid adjustments are associated with atmospheric temperature changes, i.e., semi-direct effects, except for EC-Earth accounting also for adjustments in cloud microphysics. A discussion of the rapid adjustments and the choice for the Twomey effect in ECHAM is given by Fiedler et al. (2017).

3   Results

3.1   Spread in present-day ERF

We characterise the spread in the shortwave effective radiative forcing (ERF) at the top of the atmosphere in our model ensemble for the present-day (mid-2000s). For doing so, we first calculate the multi-model mean as a reference value. The all-sky top-of-atmosphere ERF for the entire multi-model, multi-member ensemble is $-0.59\,\mathrm{Wm^{-2}}$ with an interannual standard deviation of $0.3\,\mathrm{Wm^{-2}}$, corresponding to a relative variability of roughly 50%. The interannual variability in ERF is illustrated by Gaussian distributions fitted to the frequency histogram in Figure 4a. The entire ~~possible~~ range in annual ERFs from the models including interannual variability is $-1.5\,\mathrm{Wm^{-2}}$ to $+0.5\,\mathrm{Wm^{-2}}$.

The all-sky ERFs from the models are $10-50\%$ less negative than the clear-sky ERF in all models, except in EC-Earth, because clouds mask the ERF of low-level anthropogenic aerosol (Table 2). That masking by clouds is most pronounced in HadGEM3. In EC-Earth, the all-sky ERF is more negative than in clear-sky because EC-Earth includes cloud lifetime effects of anthropogenic aerosols, thus simulating a stronger F_{aci} than all other participating models. The long-term averaged ERFs of ECHAM and ECHAM-HAM are similar, despite ECHAM using a prescribed climatology of τ_p and ECHAM-HAM simulating τ_p interactively (Section 2.1). This similarity suggests that the sub-monthly variability of natural aerosol does not substantially affect the mean ERF of anthropogenic aerosol, as long as F_{aci} is treated consistently in the two models. Using different parameterizations for F_{aci} can change this result because of non-linear processes. The magnitude of F_{aci}, however, remains uncertain (Bellouin et al., in prep.). One contributing uncertainty is the poor quantitative understanding of the pre-industrial aerosols (e.g., Carslaw et al., 2013).

The multi-model spread in the ensemble mean all-sky ERF of individual models is rather small, with a range of $-0.40\,\mathrm{Wm^{-2}}$ to $-0.9\,\mathrm{Wm^{-2}}$, compared to the internal variability of the entire multi-model ensemble (Fig. 4a). This multi-model spread corresponds to a range of deviations ~~to~~ from the multi-model mean of just $-0.31\,\mathrm{Wm^{-2}}$ to $+0.19\,\mathrm{Wm^{-2}}$ and is even smaller when the ERF of EC-Earth, which includes cloud-lifetime effects, is excluded. ~~The rather small multi-model spread is noteworthy because the models differ in all other aspects than the treatment of the anthropogenic aerosol, especially in the representation of clouds (see~~ One could expect less model diversity in all-sky ERF from our study than from previous inter-comparison projects (e.g., Myhre et al., 2013; Shindell et al., 2013), because we prescribe the same aerosol optical properties and the associated change in cloud droplet numbers. However, our model diversity in clear-sky ERF is smaller than for our all-sky ERF (Table 2). This points to the influence of model differences in representing clouds (Appendix B) ~~.~~ on the all-sky ERF. Our results therefore indicate that model differences in meteorological parameters contribute to the model diversity in all-sky ERF. This is also the case for the ERF uncertainty in a complex aerosol-climate model (Regayre et al., 2018).

The large interannual variability implies that it is essential to estimate ERF of individual models from a sufficiently large number of simulated years to quantify model differences in ERF. Otherwise the modelled ERF estimates may not be representative of the long-term average. This could be done either from sufficiently long simulations with annually repeating aerosol or a sufficiently large ensemble of simulations with transient changes. Given the similar year-to-year variability in ERF in the

models, the confidence estimates from ECHAM (Fiedler et al., 2017) are a reasonable approximation for the whole ensemble of models in the present study.

3.2 Regional contributions to ERF

The distributions of ERF for 2005 are shown as ensemble averages in Figure 5 and for each model in Figure 6. East Asia is the largest contributor to globally-averaged ERF, as expected from the regional maximum in τ_a prescribed there (Figure 5b). The mean pattern of regional contributions to ERF is in general similar in the models but differences in its magnitude and detectability appear in some regions. For example, the contributions to the global ERF modelled over central Africa range from positive to negative, averaging to a small value in this region (Fig. 5).

Another interesting example for where regional contributions to globally-averaged ERF differ is the North Atlantic. In this region, the variability of the multi-model ensemble is relatively large, $3-6$ Wm^{-2} (Figure 4b), but the small multi-model mean radiative effects are nevertheless detectable (Figure 5), although ECHAM and HadGEM by themselves have regional signals over the North Atlantic that are not statistically significant.

Taken together, the size of year-to-year variability and regional model differences in contributions to the global ERF imply that an ensemble of simulations with more than one model, as done here, is needed for constraining the radiative effect of anthropogenic aerosol regionally. The spread in modelled regional contributions to ERF is typically smaller than the differences associated with natural variability in the model ensemble (Figure 4b$-$c). Irrespectively whether we compute the regional standard deviations for the aerosol pattern of the mid-1970s or the mid-2000s, the pattern and strength of the regional natural variability in contributions to ERF is robust (not shown). In regions where the anthropogenic aerosol burden was relatively large in 2005, like East Asia, the models disagree on the magnitude of the regional contributions to ERF (Figure 4c), which means that even for a relatively large anthropogenic aerosol optical depth, natural variability of the atmosphere remains a hurdle against constraining the regional radiative effect.

3.3 Contributions from RF and adjustments

The modelled ERF is decomposed into the contributions of rapid adjustments and RF by diagnosing the latter from double calls to the radiation scheme in the models with this functionality (Figure 5). The RF is considerably less variable from year to year than ERF. Moreover, RF clearly dominates the ERF magnitude in all models that use η_N in the radiation transfer calculation (Table 2). Remember that these models consider F_{aci} from the Twomey effect only. The net contribution of rapid adjustments to the global mean ERF ranges from 0.03 Wm^{-2} in NorESM to 0.2 Wm^{-2} in ECHAM-HAM at TOA, and acts to weaken the forcing magnitude. The positive net contribution from adjustments is consistent with buffering of perturbations by atmospheric processes.

We compare the climate model estimates of RF with the results of the offline radiation-transfer calculations described in Section 2.1. The offline estimates of the all-sky RF with MACv2-SP (Offline-v1-SP and Offline-v2-SP) are in close agreement with the RF of the climate models that represent F_{aci} in form of the Twomey effect. This agreement is remarkable since the

aerosol-climate models and the offline model differ in many aspects, including again the representation of clouds (see Appendix B).

3.4 Uncertainties in RF

The offline radiation-transfer model is used to assess the role of uncertainty in τ_p and τ_a in total RF uncertainty. The aerosol
classification of MACv2 (Offline-v2) is used as an alternative representation to MACv1 (Offline-v1). MACv2 classifies more ambiguous cases of fine-mode aerosol as anthropogenic than MACv2-SP. These cases primarily occur in remote uninhabited regions such as the Southern Ocean and the Saharan desert. These regions are poorly captured by the ground-based observation network so there the MACv2 product primarily uses global model results for separating anthropogenic from natural aerosols. Classifying additional fine-mode aerosol as anthropogenic increases the all-sky RF to -1.1 Wm^{-2}, which primarily arises due
to stronger F_{aci} in MACv2. Ambiguous aerosol classifications, which occur especially in regions with a generally low aerosol burden, and a poor observational coverage are therefore causes of uncertainty in present-day RF, with the RF getting more negative with increasing τ_a.

An even more negative RF is obtained from the offline model, namely an all-sky RF of -1.4 Wm^{-2}, when both a larger anthropogenic fraction and the lower background burden of 1750 from MACv1 (Offline-v1) is used. Note that the clear-sky
RFs from the offline estimates and the climate models are in good agreement, such that most of the uncertainty stems from the uncertain magnitude of F_{aci}. This underlines again the importance of the aerosol background for quantifying the cloudy-sky contribution to all-sky RF in agreement with previous studies (Carslaw et al., 2013; Fiedler et al., 2017). Quantitative changes in natural aerosol burden between the pre-industrial and present-day remain poorly constrained. Since the aerosol of 1750 or 1850 has not been observed, using the present-day natural aerosol as a background could yield a better comparability
of observational and model estimates in future inter-comparison studies. By prescribing both the same natural and anthropogenic aerosol across different models, differences in the radiative effects of the aerosol can be attributed to model errors in representing meteorological processes and radiative transfer.

3.5 Impact of spatial change of pollution

Although the global mean τ_a is similar for 1975 and 2005, the anthropogenic pollution covers very different regions, with the
largest maxima in Europe and the U.S. during the mid-1970s and in East Asia during the mid-2000s. The regional differences in clouds, insolation and surface albedo can contribute to changes in radiative effects that can result in a different global ERF. For instance, Figures A1-A3 show the spatial patterns of cloud properties and the surface albedo illustrating both the regional differences and the model diversity for their representation (see Appendix B). The different spatial distributions of τ_a clearly change the pattern of the radiative forcing (Figure 7). As expected, the maxima in regional contributions to RF and ERF occur
over Europe and the U.S. in the mid-1970s and over East Asia for the mid-2000s.

Despite those regional differences in radiative effects and the inter-model spread in ensemble-averaged global mean RF and ERF, the spatial pattern of τ_a has little impact on the global mean RF and ERF in each of the participating models. The

model ensemble mean changes from $-0.54\,\mathrm{Wm}^{-2}$ for the mid-1970s to $-0.59\,\mathrm{Wm}^{-2}$ for the mid-2000s. The mean monthly contributions to RF are also similar for both τ_a patterns, irrespectively which model we choose (not shown).

The ensemble-averaged change in ERF is small relative to the natural interannual variability in modelled ERFs (Figure 8). Indeed, contrasting one-year estimates from the two τ_a patterns results in a large spread in ERF changes ranging from decreases to increases in ERF with τ_a patterns (Figure 8c−d). This result is in agreement with previous findings based on ECHAM only (Fiedler et al., 2017). The result underlines again the importance of using a large number of simulated years for determining changes in ERF from free-running climate models. Moreover, it provides evidence that the global mean ERF does not strongly depend on the regional distribution of anthropogenic aerosol in the northern hemisphere.

The cloudy- and clear-sky contributions to the all-sky efficiency of the ERF, in other words the ratio of ERF to τ_a, helps to better understand why the two τ_a patterns yield similar ERFs. All-sky efficiency is the sum of contributions from cloudy and clear-sky conditions:

$$\frac{\mathrm{ERF}_{\mathrm{all}}}{\tau_a} = f\frac{\mathrm{ERF}_{\mathrm{cloudy}}}{\tau_a} + (1-f)\frac{\mathrm{ERF}_{\mathrm{clear}}}{\tau_a}, \tag{1}$$

where f is the total cloud fraction, and $\mathrm{ERF}_{\mathrm{cloudy}}$ and $\mathrm{ERF}_{\mathrm{clear}}$ the ERF in cloudy and clear-sky conditions, respectively.

Figure 9 shows the regional distribution from the multi-model ensemble average of the terms of Equation 1. The all-sky efficiency often increases with increasing distance to major pollution sources because of the decreasing background aerosol, up to $-100\,\mathrm{Wm}^{-2}$ per unit τ_a. These all-sky efficiencies are primarily explained by the cloudy-sky contributions. ~~A~~ Large efficiencies occur typically in remote areas including some regions at the edges of τ_a plumes (Fig. 9). No clear saturation of F_{aci} ~~towards the~~ is evident at all edges of the τ_a plumes~~is not evident and~~. Also the spatial distribution of both the all- and cloudy-sky efficiency is rather inhomogeneous. The inhomogeneity contrasts with the clear-sky efficiency, which has much smaller spatial variability.

Averaged globally, all-sky forcing efficiencies for the two aerosol patterns are similar at $-26\,\mathrm{Wm}^{-2}$ per unit τ_a. The regional all-sky ERF efficiencies, however, change between the mid-1970s and mid-2000s (Fig. 9). This change is almost exclusively explained by the cloudy-sky contribution to the ERF efficiency, reflecting the regional change in η_N from the mid-1970s to mid-2000s. The strong change in the cloudy-sky contribution is in strong contrast to the relatively minor changes in the clear-sky contributions. Differences in regional efficiencies of anthropogenic aerosol effects on clouds thus balance in the global mean and result in similar global ERFs for the mid-1970s and mid-2000s.

Of all models, NorESM and EC-Earth have the strongest ~~effective radiative forcing~~ ERF efficiencies around -30 and $-40\,\mathrm{Wm}^{-2}$ per unit τ_a, respectively, i.e., the same aerosol perturbation in these two models is much more efficient in inducing effective radiative effects than in the other models, consistent with the more negative ERFs (Fig. 8). In EC-Earth, the more negative ERF arises from also perturbing the cloud microphysics with η_N. In NorESM, the more negative ERF arises from a strong negative RF and a small net contribution from adjustments.

4 Conclusions

We assess the radiative effects of anthropogenic aerosol in ensembles of simulations from five state-of-the-art aerosol climate models, prescribing identical anthropogenic aerosol properties of reduced complexity. Each of the participating models uses annually repeating patterns of anthropogenic aerosol for obtaining 180 years of radiative forcing estimates. The multi-model multi-ensemble present-day all-sky short-wave effective radiative forcing (ERF) at the top of atmosphere is -0.59 Wm^{-2}. The year-to-year standard deviations of around 0.3 Wm^{-2} in the models imply a typical year-to-year variability of 50%, reflecting a strong contribution of model internal variability to ERF. We therefore recommend caution for the use of ERF estimates based on single years, as in the standard AeroCom protocol with varying reference years. These are likely affected by model-internal variability such that an apparent ERF spread is not associated with systematic model differences alone. Indeed such studies have shown a substantial spread in ERF estimates (e.g., Shindell et al., 2013), comparable to the magnitude of the model internal variability quantified in the present work.

We further recommend that model-based assessments of ERF in the future ensure to eliminate the effects of internal variability, either by averaging over longer time periods from single transient climate simulations or from averaging across several ensemble members for shorter time periods. For instance, the protocol of RFMIP requests thirty-year averages for estimating the present-day ERF and three-member ensembles with ten-year averages for diagnosing decadal changes in ERF (Pincus et al., 2016). The precision of the estimate can be tested by using confidence estimates (e.g., Fiedler et al., 2017). Note that natural variability is equally an issue in observations. Ensembles of simulations should therefore be used for constraining ERF with the historical record of observations. The interannual variability in ERF, and hence the number of years needed to estimate ERF, could be different in nudged model simulations (Zhang et al., 2014). However, nudging a model simulation with re-analysis data can change the climatology and interfere with the rapid adjustments. The resulting ERFs from a nudged simulation are therefore likely different compared with free-running model simulations. The interference of nudging with adjustments deserves closer attention in future research.

In our study, we obtain an ERF spread of -0.9 to -0.4 Wm^{-2} associated with systematic model differences (Fig. 10). This estimate is not affected by model-internal variability, is based on identical anthropogenic aerosol optical properties and makes use of a consistent perturbation of the cloud droplet number concentrations associated with anthropogenic aerosol. The model with the most negative ERF accounts also for changes in cloud microphysics associated with anthropogenic aerosol, whereas the other participating models account for the Twomey effect only. Based on our model spread, we conclude that models with a strongly negative ERF have particularly strong contributions from anthropogenic aerosol effects on clouds.

Our results highlight that the participating models consistently show little change in the global ERF of anthropogenic aerosol between the mid-1970s and mid-2000s, despite the substantially different location of anthropogenic pollution maxima and the model diversity in their ERF magnitude relative to the pre-industrial. Model internal variability, however, produces ERF changes of different signs and magnitude between the two periods. This result gives further evidence that model-internal variability has not been sufficiently considered in past model studies estimating the ERF difference associated with the mid-1970s to mid-2000s change in anthropogenic aerosol, as previously suggested based on ECHAM alone (Fiedler et al., 2017).

The small change in global ERF stems from similar global forcing efficiencies of anthropogenic aerosol in the two periods. These are primarily explained by globally compensating differences in regional cloudy-sky contributions to the ERF efficiency. Assuming stronger aerosol-cloud interactions can cause a larger change in ERF from the mid-1970s to mid-2000s, based on simulations with ECHAM (Fiedler et al., 2017). The forcing from aerosol-cloud interaction is a subject of ongoing discussion

5 and research (Bellouin et al., in prep.). Given our multi-model spread in absolute ERF relative to the pre-industrial, inter-comparing the relative ERF changes between observable periods might provide a better test for a model to represent transient climate changes. Our future work will focus on inter-comparing modelled ERF changes associated with other aerosol patterns. One such endeavour is the usage of MACv2-SP in model simulations in the framework of CMIP6 (e.g., Pincus et al., 2016; Fiedler et al., 2018).

10 *Data availability.* The model data of this study will be available on the AeroCom community's data server. Additionally, the model data is archived by the Max Planck Institute for Meteorology and can be made accessible by contacting publications@mpimet.mpg.de.

Appendix A: Model physics packages

ECHAM6.3 is the latest version of the atmosphere component of the Earth system model MPI-ESM1.2 of MPI-M, which participates in CMIP6 (Mauritsen et al., 2019). ECHAM6.3 is a global hydrostatic model and includes parameterisations of

15 sub-grid scale physical processes. The atmospheric radiative transfer is parameterised with the PSrad scheme using the Rapid Radiative Transfer Model for general circulation models (RRTMG, Pincus and Stevens, 2013). Surface properties, trace gas concentrations, and natural aerosols are prescribed by climatological data sets. A major change in MPI-ESM1.2 (Mauritsen et al., 2019) compared to previous model versions is the implementation of MACv2-SP (Fiedler et al., 2017; Stevens et al., 2017).

20 The global aerosol-climate model ECHAM6.3-HAM2.3 is an updated version of the model described by Tegen et al. (2018) and Neubauer et al. (submitted to ACPD). Revisions made in ECHAM6.3-HAM2.3 relate to the atmospheric model and the description of sea-salt emissions, which have been made dependent on the sea-surface temperature. The model uses ECHAM6.3, but is coupled to the aerosol module HAM (Stier et al., 2005; Zhang et al., 2012). An important difference in the atmospheric components is that ECHAM6.3 uses a single-moment cloud microphysics parameterisation, while ECHAM6.3-HAM2.3 has

25 a two-moment stratiform cloud scheme (Lohmann and Hoose, 2009) for representing the activation of aerosols as cloud condensation nuclei and ice nuclei in mixed phase clouds. Emission schemes for sea salt (Long et al., 2011; Sofiev et al., 2011), desert dust (Tegen et al., 2002; Cheng et al., 2008), and oceanic dimethyl sulphide (DMS, Nightingale et al., 2000) are run online. Emission of all other aerosol species are prescribed from external input files (Stier et al., 2005; Lamarque et al., 2010). In the configuration used in this study, we prescribe the pre-industrial background of aerosol components from HAM that are

30 not simulated online. These, in combination with the online-computed natural aerosol emissions, are the only aerosols seen by the two-moment cloud microphysics parameterisation in this study.

EC-Earth (Hazeleger et al., 2010; Döscher et al., in prep.) uses the Integrated Forecasting System (IFS) of the European Centre for Medium-range Weather Forecasts (ECMWF) as its atmosphere component. The latest generation of the model, EC-Earth3, is based on the ECMWF seasonal prediction system 4 with IFS cycle 36r4. The radiation scheme is based on the Rapid Radiative Transfer Model (Mlawer and Clough, 1998; Iacono et al., 2008) with 14 bands in the shortwave and 16 bands in the longwave spectrum, and uses the Monte-Carlo Independent Column Approximation (McICA) approach (Pincus and Morcrette, 2003). Many new features have been added to IFS by the EC-Earth consortium. The pre-industrial tropospheric aerosol climatology that is used in combination with MACv2-SP, has been constructed from a simulation with the TM5 aerosol-chemistry model (Huijnen et al., 2010; van Noije et al., 2014), driven by meteorological data from ERA-Interim for the early 1980s. This simulation used CMIP6 emissions of aerosol and precursor gases for 1850, and provides the monthly mean aerosol mass and number concentrations as well as the aerosol optical properties. Stratospheric aerosols are prescribed using the CMIP6 data set of radiative properties. Aerosol-cloud interactions are implemented only for liquid phase, stratiform clouds. The cloud droplet number concentration, N, is diagnosed using the activation scheme of Abdul-Razzak and Ghan (2000) and is here modified by η_N from MACv2-SP. Cloud microphysics depends on N through autoconversion of cloud droplets to rain. The model used in this study is EC-Earth version 3.2.3. It is close to the CMIP6 version described by Döscher et al. (in prep.), but does not include the latest revisions that were introduced after the simulations for this study were started. Most relevant for this study is that in the CMIP6 version the pre-industrial aerosol climatology has been updated, by changing the parameterization of the production of sea spray in the underlying TM5 model. Specifically, the whitecap coverage has been made dependent on sea-surface temperature, while its power-law dependence on the 10m-wind speed has been changed from the W10 expression proposed by Salisbury et al. (2013) to the expression proposed by Monahan and Muircheartaigh (1980). The main effect of this revision is an increase in aerosol and cloud droplet number concentrations over the Southern Ocean.

Simulations with the Hadley Centre Global Environment Model (HadGEM) use a modified version of the HadGEM3 Global Atmosphere 7.0 climate model configuration (Walters et al., 2017). HadGEM3 normally uses the Global Model of Aerosol Processes (GLOMAP, Mann et al., 2010) to simulate aerosol mass and number, and interactions of aerosols with radiation, clouds and atmospheric chemistry. That scheme is here replaced with prescriptions of the three-dimensional distributions of aerosol extinction and absorption coefficients averaged over HadGEM's 6 shortwave and 9 longwave wavebands, waveband-averaged aerosol asymmetry, and N. Those prescriptions are made of three components. First, pre-industrial aerosol and N distributions are taken from a HadGEM3/GLOMAP simulation using CMIP6 emission datasets for the year 1850. Second, stratospheric aerosols are taken from the CMIP6 climatologies for the year 1850. Prescribed N are used in the calculation of cloud albedo (Jones et al., 2001) and autoconversion rates (Khairoutdinov and Kogan, 2000), although the latter do not see the MACv2-SP N scalings, ensuring that anthropogenic aerosols do not exert a secondary indirect effect in the present study. HadGEM3 uses the Prognostic Cloud fraction and Prognostic Condensate scheme (PC2, Wilson et al., 2008) that simulates the mass-mixing ratios of water vapour, cloud liquid and ice, as well as the fractional cover of liquid, ice, and mixed-phase clouds.

The Norwegian Earth System Model (NorESM Bentsen et al., 2013; Iversen et al., 2013; Kirkevåg et al., 2013) uses the atmospheric component of the Oslo version of the Community Atmosphere Model (CAM4-Oslo), which differs from the original CAM4 (Neale et al., 2013) through the modified treatment of aerosols and their interaction with clouds (Kirkevåg

et al., 2013). The model has a finite-volume dynamical core and the original version 4 of the Community Land Model (CLM4) of CCSM4 (Lawrence et al., 2011). NorESM uses the CAM-RT radiation scheme by Collins et al. (2006). Like ECHAM-HAM and ECHAM, NorESM sets all background aerosol emission to pre-industrial levels representative of 1850. These background conditions include sulphate from tropospheric volcanoes and from DMS, as well as organic matter from land

5 and ocean biogenic processes, mineral dust and sea salt. Sea salt emissions are parameterised as a function of wind speed and temperature (Struthers et al., 2011), while other pre-industrial aerosol emissions are prescribed following Kirkevåg et al. (2013). These are, in the case of NorESM, sulphate, organic matter and BC aerosols originating from fossil fuel emissions and biomass burning (Lamarque et al., 2010).

Appendix B: Model diversity in cloud properties and surface albedo

10 The model diversity in RF and ERF is larger when cloudy skies are considered. We therefore assess the model diversity in cloud properties and compare the model climatologies calculated from the simulations for the mid-2000s against observational climatologies from satellite products, listed in Table A1. The observational products provide an orientation for realistic values, although satellite retrievals also have caveats (e.g., Grosvenor et al., 2018). Moreover, we document the here-used surface albedos for illustrating both the regional differences and the model diversity.

15 **B1   Macroscopic cloud properties**

We first assess the cloud shortwave radiative effect at the top of the atmosphere (F_{cld}), thus the cloud effect on the planetary albedo. The multi-annual global mean F_{cld} for $2001-2010$ from CERES Ed. 4 is -45.8 Wm^{-2}, i.e., less negative than in most models (Table A2). This behaviour indicates a tendency of the models to have too reflective clouds consistent with other model evaluations (Nam et al., 2012; Crueger et al., 2018, Lohmann and Neubauer, submitted). The spatial patterns of modelled F_{cld}

20 are generally similar, but regionally the differences can be more distinct (Figure A1).

To better characterise the model diversity in clouds, we compare the simulated total cloud cover (f) and liquid water path (l_{cld}) to satellite climatologies from ISCCP and MAC-LWP, respectively (Table A1). Most models underestimate both f and l_{cld} over the oceans compared to the satellite retrievals, but having too few clouds does not necessarily imply too small amount of liquid or vice versa (Table A2). The spatial patterns (Figure A1) show a tendency of the models for underestimating f in

25 the stratocumulus decks in the Southeastern regions of the Pacific and Atlantic Ocean, where aerosol-cloud interactions are thought to be important. The models, however, disagree on the values for f and l_{cld} in those regions. Moreover, the models show a large diversity in l_{cld} in the extra-tropical storm tracks. NorESM shows the largest maximum l_{cld} exceeding 200 gm^{-2}. Our findings for l_{cld} are consistent with a similar regional comparison between HadGEM and CAM (Malavelle et al., 2017), the latter of which has a similar atmospheric component as NorESM (see Appendix A).

13

B2 Cloud microphysical properties

The reported differences in macroscopic cloud properties among the models raise the question how different the cloud droplet number concentrations (N) are. We find that the models show large diversity in the pattern of N for present-day conditions as shown in Figure A2. Note that we show the mean in-cloud droplet number concentration, which means that regions without clouds are not included when averaging N. It is noteworthy that in the models N is calculated for stratiform cloud types, but can additionally include detrained droplets from anvils of deep convection. The spatial pattern of N in ECHAM is not shown due to the simplistic treatment in the model. ECHAM employs statically prescribed values for N, which are constant with height below 800 hPa and exponentially decrease aloft. The near-surface values in ECHAM are N=80 cm^{-3} over ocean and N=180 cm^{-3} elsewhere (not shown), and are multiplied with η_N from MACv2-SP like in the other models.

Compared to the satellite product, the models typically underestimate N, e.g., in the stratocumulus decks, where also f is underestimated. It remains an open question how much of the quantitative differences between the models and the satellite product is due to differences in the methods for diagnosing N in the satellite retrievals and the models, but it is unlikely that the methods solely explain the diversity in the patterns of N. It is interesting that, despite these quantitative differences in N, the spatial pattern of F_{cld} compares reasonably well to observations (Figure A1), which might be a consequence of compensating differences from tuning the radiation balance at the top of the atmosphere. For instance, the behaviour of NorESM points to too much shortwave reflectivity by too thick clouds that overcompensate the missing reflection due to underestimated cloud cover.

B3 Surface albedo

An additional influence on the radiative forcing of anthropogenic aerosol is the surface reflectivity for shortwave radiation. We therefore document the surface albedo for shortwave radiation from the participating models and the satellite product used in the offline radiative transfer calculations of this study. In the global mean, the models and the satellite product are very similar, with a surface albedo of $14-16\%$. However, the spatial distributions in Figure A3 indicate differences. The typical difference between less reflective ocean surfaces compared to land regions is apparent. Moreover, the analysis reveals diversity in the regional surface albedos of the participating models, typically related to areas affected by snow cover. Since such diversity in the surface albedo was already previously reported for aerosol-climate models with implications for the aerosol radiative forcing (e.g., Stier et al., 2007), future efforts are still needed for constraining the surface albedo in climate models.

15

References

Abdul-Razzak, H. and Ghan, S. J.: A parameterization of aerosol activation: 2. Multiple aerosol types, J. Geophys. Res., pp. 6837–6844, doi:10.1029/1999JD901161, 2000.

Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., Boucher, O., Carslaw, K., Christensen, M., Daniau, A.-L., Dufrsene, J.-L., Feingold, G., Fiedler, S., Forster, P., Gettleman, A., Gryspeerdt, EA., Haywood, J., Malavelle, F., Lohmann, U., Mauritsen, T., McCoy, D., Myhre, G., Mülmenstädt, J., Neubauer, D., Possner, A., Quaas, J., Rugenstein, M., Sato, Y., Schulz, M., Schwarz, S.E., Sourdeval, O., Storelvmo, T., Toll, V., Watson-Parris Winker, D., and et al. Stevens, B.: Bounding aerosol radiative forcing , on Climate, Reviews in Geophysics, in prep.

Bennartz, R. and Rausch, J.: Global and regional estimates of warm cloud droplet number concentration based on 13 years of AQUA-MODIS observations, Atmospheric Chemistry and Physics, 17, 9815–9836, doi:10.5194/acp-17-9815-2017, https://www.atmos-chem-phys.net/17/9815/2017/, 2017.

Bentsen, M., Bethke, I., Debernard, J. B., Iversen, T., Kirkevåg, A., Seland, Ø., Drange, H., Roelandt, C., Seierstad, I. A., Hoose, C., and Kristjánsson, J. E.: The Norwegian Earth System Model, NorESM1-M – Part 1: Description and basic evaluation of the physical climate, Geoscientific Model Development, 6, 687–720, doi:10.5194/gmd-6-687-2013, https://www.geosci-model-dev.net/6/687/2013/, 2013.

Carslaw, K., Lee, L., Reddington, C., Pringle, K., Rap, A., Forster, P., Mann, G., Spracklen, D., Woodhouse, M., Regayre, L., et al.: Large contribution of natural aerosols to uncertainty in indirect forcing, Nature, 503, 67–71, 2013.

Cheng, T., Peng, Y., Feichter, J., and Tegen, I.: An improvement on the dust emission scheme in the global aerosol-climate model ECHAM5-HAM, Atmospheric Chemistry and Physics, 8, 1105–1117, doi:10.5194/acp-8-1105-2008, https://www.atmos-chem-phys.net/8/1105/2008/, 2008.

Collins, W. D., Rasch, P. J., Boville, B. A., Hack, J. J., McCaa, J. R., Williamson, D. L., Briegleb, B. P., Bitz, C. M., Lin, S.-J., and Zhang, M.: The Formulation and Atmospheric Simulation of the Community Atmosphere Model Version 3 (CAM3), Journal of Climate, 19, 2144–2161, doi:10.1175/JCLI3760.1, https://doi.org/10.1175/JCLI3760.1, 2006.

Crueger, T., Giorgetta, M. A., Brokopf, R., Esch, M., Fiedler, S., Hohenegger, C., Kornblueh, L., Mauritsen, T., Nam, C., Naumann, A. K., Peters, K., Rast, S., Roeckner, E., Sakradzija, M., Schmidt, H., Vial, J., Vogel, R., and Stevens, B.: ICON-A, the atmosphere component of the ICON Earth System Model. Part II: Model evaluation, Journal of Advances in Modeling Earth Systems, 0, doi:10.1029/2017MS001233, 2018.

Dentener, F., Kinne, S., and Bond, T.: Emissions of primary aerosol and precursor gases in the years 2000 and 1750 prescribed data-sets for AeroCom, Atmos. Chem. Phys., 6, 4321–4344, 2006.

Döscher, R. et al.: The community Earth system model EC-Earth for collaborative climate research, in prep.

Elsaesser, G. S., O'Dell, C. W., Lebsock, M. D., Teixeira, J.: Multisensor Advanced Climatology Mean Liquid Water Path L3 Monthly 1 degree x 1 degree V1, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: March 2018, doi:10.5067/MEASURES/MACLWPM, 2016.

Elsaesser, G. S., O'Dell, C. W., Lebsock, M. D., Bennartz, R., Greenwald, T. J., and Wentz, F. J.: The Multisensor Advanced Climatology of Liquid Water Path (MAC-LWP), Journal of Climate, 30, 10 193–10 210, doi:10.1175/JCLI-D-16-0902.1, 2017.

Evan, A. T., Flamant, C., Fiedler, S., and Doherty, O.: An analysis of aeolian dust in climate models, Geophys. Res. Let., doi:10.1002/2014GL060545, 2014.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958, doi:10.5194/gmd-9-1937-2016, 2016.

Fiedler, S., Knippertz, P., Woodward, S., Martin, G. M., Bellouin, N., Ross, A. N., Heinold, B., Schepanski, K., Birch, C. E., and Tegen, I.: A process-based evaluation of dust-emitting winds in the CMIP5 simulation of HadGEM2-ES, Climate Dynamics, 46, 1107–1130, doi:10.1007/s00382-015-2635-9, 2016.

Fiedler, S., Stevens, B., and Mauritsen, T.: On the sensitivity of anthropogenic aerosol forcing to model-internal variability and parameterizing a Twomey effect, J. Adv. Model. Earth Syst., pp. n/a–n/a, doi:10.1002/2017MS000932, 2017.

Fiedler, S., Stevens, B., Gidden, M., Smith, S. J., Riahi, K., and van Vuuren, D.: First forcing estimates from the future CMIP6 scenarios of anthropogenic aerosol optical properties and an associated Twomey effect, Geosci. Model Dev., 12, 989–1007, doi:10.5194/gmd-12-989-2019, 2019.

Grosvenor, D. P., Sourdeval, O., Zuidema, P., Ackerman, A., Alexandrov, M. D., Bennartz, R., Boers, R., Cairns, B., Chiu, J. C., Christensen, M., Deneke, H., Diamond, M., Feingold, G., Fridlind, A., Hünerbein, A., Knist, C., Kollias, P., Marshak, A., McCoy, D., Merk, D., Painemal, D., Rausch, J., Rosenfeld, D., Russchenberg, H., Seifert, P., Sinclair, K., Stier, P., van Diedenhoven, B., Wendisch, M., Werner, F., Wood, R., Zhang, Z., and Quaas, J.: Remote Sensing of Droplet Number Concentration in Warm Clouds: A Review of the Current State of Knowledge and Perspectives, Reviews of Geophysics, 56, 409–453, doi:10.1029/2017RG000593, 2018.

Halthore, R. N., Crisp, D., Schwartz, S. E., Anderson, G. P., Berk, A., Bonnel, B., Boucher, O., Chang, F.-L., Chou, M.-D., Clothiaux, E. E., Dubuisson, P., Fomin, B., Fouquart, Y., Freidenreich, S., Gautier, C., Kato, S., Laszlo, I., Li, Z., Mather, J. H., Plana-Fattori, A., Ramaswamy, V., Ricchiazzi, P., Shiren, Y., Trishchenko, A., and Wiscombe, W.: Intercomparison of shortwave radiative transfer codes and measurements, Journal of Geophysical Research: Atmospheres, 110, n/a–n/a, doi:10.1029/2004JD005293, http://dx.doi.org/10.1029/2004JD005293, d11206, 2005.

Hazeleger, W., Severijns, C., Semmler, T., Stefanescu, S., Yang, S., Wang, X., Wyser, K., Dutra, E., Baldasano, J. M., Bintanja, R., Bougeault, P., Caballero, R., Ekman, A. M. L., Christensen, J. H., van den Hurk, B., Jimenez, P., Jones, C., Kalberg, P., Koenigk, T., McGrath, R., Miranda, P., van Noije, T., Palmer, T., Parodi, J. A., Schmith, T., Selten, F., Storelvmo, T., Sterl, A., Tapamo, H., Vancoppenolle, M., Viterbo, P., and Willen, U.: EC-Earth, Bulletin of the American Meteorological Society, 91, 1357–1364, doi:10.1175/2010BAMS2877.1, 2010.

Huijnen, V., Williams, J., van Weele, M., van Noije, T., Krol, M., Dentener, F., Segers, A., Houweling, S., Peters, W., de Laat, J., Boersma, F., Bergamaschi, P., van Velthoven, P., Le Sager, P., Eskes, H., Alkemade, F., Scheele, R., Nédélec, P., and Pätz, H.-W.: The global chemistry transport model TM5: description and evaluation of the tropospheric chemistry version 3.0, Geoscientific Model Development, 3, 445–473, doi:10.5194/gmd-3-445-2010, 2010.

Huneeus, N., Basart, S., Fiedler, S., Morcrette, J.-J., Benedetti, A., Mulcahy, J., Terradellas, E., Pérez García-Pando, C., Pejanovic, G., Nickovic, S., Arsenovic, P., Schulz, M., Cuevas, E., Baldasano, J. M., Pey, J., Remy, S., and Cvetkovic, B.: Forecasting the northern African dust outbreak towards Europe in April 2011: a model intercomparison, Atmospheric Chemistry and Physics, 16, 4967–4986, doi:10.5194/acp-16-4967-2016, 2016.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, Journal of Geophysical Research: Atmospheres, 113, doi:10.1029/2008JD009944, 2008.

Iversen, T., Bentsen, M., Bethke, I., Debernard, J. B., Kirkevåg, A., Seland, Ø., Drange, H., Kristjansson, J. E., Medhaug, I., Sand, M., and Seierstad, I. A.: The Norwegian Earth System Model, NorESM1-M – Part 2: Climate response and scenario projections, Geoscientific Model Development, 6, 389–415, doi:10.5194/gmd-6-389-2013, https://www.geosci-model-dev.net/6/389/2013/, 2013.

Jones, A., Roberts, D. L., Woodage, M. L., and Johnson, C. E.: Indirect sulphate forcing in a climate model with an interactive sulphur cycle, J. Geophys. Res., pp. 20,293–20,310, 2001.

Khairoutdinov, M. F. and Kogan, Y. L.: A new cloud physics parameterization in a large-eddy simulation model of marine stratocumulus, Mon. Wea. Rev., p. 229–243, 2000.

Kinne, S.: The Max-Planck Aerosol Climatology, version 2, in review.

Kinne, S., Schulz, M., Textor, C., Guibert, S., Balkanski, Y., Bauer, S. E., Berntsen, T., Berglen, T. F., Boucher, O., Chin, M., Collins, W., Dentener, F., Diehl, T., Easter, R., Feichter, J., Fillmore, D., Ghan, S., Ginoux, P., Gong, S., Grini, A., Hendricks, J., Herzog, M., Horowitz, L., Isaksen, I., Iversen, T., Kirkevåg, A., Kloster, S., Koch, D., Kristjansson, J. E., Krol, M., Lauer, A., Lamarque, J. F., Lesins, G., Liu, X., Lohmann, U., Montanaro, V., Myhre, G., Penner, J., Pitari, G., Reddy, S., Seland, O., Stier, P., Takemura, T., and Tie, X.: An AeroCom initial assessment - optical properties in aerosol component modules of global models, Atmospheric Chemistry and Physics, 6, 1815–1834, doi:10.5194/acp-6-1815-2006, 2006.

Kinne, S., O'Donnel, D., Stier, P., Kloster, S., Zhang, K., Schmidt, H., Rast, S., Giorgetta, M., Eck, T. F., and Stevens, B.: MAC-v1: A new global aerosol climatology for climate studies, Journal of Advances in Modeling Earth Systems, 5, 704–740, doi:10.1002/jame.20035, 2013.

Kirkevåg, A., Iversen, T., Seland, Ø., Hoose, C., Kristjánsson, J. E., Struthers, H., Ekman, A. M. L., Ghan, S., Griesfeller, J., Nilsson, E. D., and Schulz, M.: Aerosol–climate interactions in the Norwegian Earth System Model – NorESM1-M, Geoscientific Model Development, 6, 207–244, doi:10.5194/gmd-6-207-2013, https://www.geosci-model-dev.net/6/207/2013/, 2013.

Koffi, B., Schulz, M., Bréon, F.-M., Dentener, F., Steensen, B. M., Griesfeller, J., Winker, D., Balkanski, Y., Bauer, S. E., Bellouin, N., Berntsen, T., Bian, H., Chin, M., Diehl, T., Easter, R., Ghan, S., Hauglustaine, D. A., Iversen, T., Kirkevåg, A., Liu, X., Lohmann, U., Myhre, G., Rasch, P., Seland, O., Skeie, R. B., Steenrod, S. D., Stier, P., Tackett, J., Takemura, T., Tsigaridis, K., Vuolo, M. R., Yoon, J., and Zhang, K.: Evaluation of the aerosol vertical distribution in global aerosol models through comparison against CALIOP measurements: AeroCom phase II results, Journal of Geophysical Research: Atmospheres, 121, 7254–7283, doi:10.1002/2015JD024639, http://dx.doi.org/10.1002/2015JD024639, 2015JD024639, 2016.

Lacagnina, C., Hasekamp, O. P., Bian, H., Curci, G., Myhre, G., van Noije, T., Schulz, M., Skeie, R. B., Takemura, T., and Zhang, K.: Aerosol single-scattering albedo over the global oceans: Comparing PARASOL retrievals with AERONET, OMI, and AeroCom models estimates, Journal of Geophysical Research: Atmospheres, 120, 9814–9836, doi:10.1002/2015JD023501, 2015JD023501, 2015.

Lamarque, J.-F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A., Owen, B., Schultz, M. G., Shindell, D., Smith, S. J., Stehfest, E., Van Aardenne, J., Cooper, O. R., Kainuma, M., Mahowald, N., McConnell, J. R., Naik, V., Riahi, K., and van Vuuren, D. P.: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application, Atmospheric Chemistry and Physics, 10, 7017–7039, doi:10.5194/acp-10-7017-2010, https://www.atmos-chem-phys.net/10/7017/2010/, 2010.

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X., Yang, Z.-L., Levis, S., Sakaguchi, K., Bonan, G. B., and Slater, A. G.: Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model, Journal of Advances in Modeling Earth Systems, 3, n/a–n/a, doi:10.1029/2011MS00045, http://dx.doi.org/10.1029/2011MS00045, m03001, 2011.

Loeb, N. G., Wielicki, B. A., Doelling, D. R., Smith, G. L., Keyes, D. F., Kato, S., Manalo-Smith, N., and Wong, T.: Toward Optimal Closure of the Earth's Top-of-Atmosphere Radiation Budget, J. Climate, 22, 748–766, doi:10.1175/2008JCLI2637.1, 2009.

Lohmann, U. and Ferrachat, S.: Impact of parametric uncertainties on the present-day climate and on the anthropogenic aerosol effect, Atm. Chem. Phys., 10, 11 373–11 383, doi:10.5194/acp-10-11373-2010, 2010.

Lohmann, U. and Hoose, C.: Sensitivity studies of different aerosol indirect effects in mixed-phase clouds, Atmospheric Chemistry and Physics, 9, 8917–8934, doi:10.5194/acp-9-8917-2009, https://www.atmos-chem-phys.net/9/8917/2009/, 2009.

Long, M. S., Keene, W. C., Kieber, D. J., Erickson, D. J., and Maring, H.: A sea-state based source function for size- and composition-resolved marine aerosol production, Atmospheric Chemistry and Physics, 11, 1203–1216, doi:10.5194/acp-11-1203-2011, 2011.

Mahowald, N. M. and Luo, C.: A less dusty future?, Geophysical Research Letters, 30, n/a–n/a, doi:10.1029/2003GL017880, http://dx.doi.org/10.1029/2003GL017880, 2003.

Malavelle, F. F., Haywood, J. M., Jones, A., Gettelman, A., Clarisse, L., Bauduin, S., Allan, R. P., Karset, I. H. H., Kristjánsson, J. E., Oreopoulos, L., Cho, N., Lee, D., Bellouin, N., Boucher, O., Grosvenor, D. P., Carslaw, K. S., Dhomse, S., Mann, G. W., Schmidt, A., Coe, H., Hartley, M. E., Dalvi, M., Hill, A. A., Johnson, B. T., Johnson, C. E., Knight, J. R., O'Connor, F. M., Partridge, D. G., Stier, P., Myhre, G., Platnick, S., Stephens, G. L., Takahashi, H., and Thordarson, T.: Strong constraints on aerosol–cloud interactions from volcanic eruptions, Nature, doi:10.1038/nature22974, 2017.

Mann, G. W., Carslaw, K. S., Spracklen, D. V., Ridley, D. A., Manktelow, P. T., Chipperfield, M. P., Pickering, S. J., and Johnson, C. E.: Description and evaluation of GLOMAP-mode: a modal global aerosol microphysics model for the UKCA composition-climate model, Geosci. Model Dev., 3, 519–551, 2010.

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenez de la Cuesta Otero, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D., Nyawira, S.-S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., von Storch, J.-S., Tian, F., Voigt, A., de Vrese, P., Wieners, K.-H., Wilkenskjeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM 1.2) and its response to increasing CO2, Journal of Advances in Modeling Earth Systems, doi:10.1029/2018MS001400, 2019.

Mlawer, E. and Clough, S.: Shortwave and longwave enhancements in the rapid radiative transfer model, in: Proceedings of the 7th Atmospheric Radiation Measurement (ARM) Science Team Meeting, pp. 409–413, 1998.

Monahan, E. C. and Muircheartaigh, I.: Optimal power-law description of oceanic whitecap coverage dependence on wind spee, J. Phys. Oceanogr, pp. 2094—-2099, doi:10.1175/1520-0485, 1980.

Myhre, G., Samset, B. H., Schulz, M., Balkanski, Y., Bauer, S., Berntsen, T. K., Bian, H., Bellouin, N., Chin, M., Diehl, T., Easter, R. C., Feichter, J., Ghan, S. J., Hauglustaine, D., Iversen, T., Kinne, S., Kirkevåg, A., Lamarque, J.-F., Lin, G., Liu, X., Lund, M. T., Luo, G., Ma, X., van Noije, T., Penner, J. E., Rasch, P. J., Ruiz, A., Seland, Ø., Skeie, R. B., Stier, P., Takemura, T., Tsigaridis, K., Wang, P., Wang, Z., Xu, L., Yu, H., Yu, F., Yoon, J.-H., Zhang, K., Zhang, H., and Zhou, C.: Radiative forcing of the direct aerosol effect from AeroCom Phase II simulations, Atmospheric Chemistry and Physics, 13, 1853–1877, doi:10.5194/acp-13-1853-2013, 2013.

Nam, C., Bony, S., Dufresne, J.-L., and Chepfer, H.: The 'too few, too bright' tropical low-cloud problem in CMIP5 models, Geophysical Research Letters, 39, n/a–n/a, doi:10.1029/2012GL053421, l21801, 2012.

Neale, R. B., Richter, J., Park, S., Lauritzen, P. H., Vavrus, S. J., Rasch, P. J., and Zhang, M.: The Mean Climate of the Community Atmosphere Model (CAM4) in Forced SST and Fully Coupled Experiments, Journal of Climate, 26, 5150–5168, doi:10.1175/JCLI-D-12-00236.1, https://doi.org/10.1175/JCLI-D-12-00236.1, 2013.

Neubauer, D., Ferrachat, S., Siegenthaler-Le Drian, C., Stier, P., Partridge, P. D., Tegen, I., Bey, I., Stanelle, T., Kokkola, H., and Lohmann, U.: Cloud evaluation, aerosol radiative forcing and climate sensitivity in the global aerosol-climate model ECHAM6.3- HAM2.3, submitted to ACPD.

Nightingale, P. D., Malin, G., Law, C. S., Watson, A. J., Liss, P. S., Liddicoat, M. I., Boutin, J., and Upstill-Goddard, R. C.: In situ evaluation of air-sea gas exchange parameterizations using novel conservative and volatile tracers, Global Biogeochemical Cycles, 14, 373–387, doi:10.1029/1999GB900091, http://dx.doi.org/10.1029/1999GB900091, 2000.

Pincus, R and, B. H. W. and Morcrette, J.-J.: A fast, flexible, approximate technique for computing radiative transfer in inhomogeneous cloud fields, J. Geophys. Res., p. 4376, doi:10.1029/2002JD003322, 2003.

Pincus, R. and Stevens, B.: Paths to accuracy for radiation parameterizations in atmospheric models, Journal of Advances in Modeling Earth Systems, 5, 225–233, doi:10.1002/jame.20027, 2013.

Pincus, R., Forster, P. M., and Stevens, B.: The Radiative Forcing Model Intercomparison Project (RFMIP): experimental protocol for CMIP6, Geoscientific Model Development, 9, 3447–3460, doi:10.5194/gmd-9-3447-2016, 2016.

Quaas, J., Ming, Y., Menon, S., Takemura, T., Wang, M., Penner, J. E., Gettelman, A., Lohmann, U., Bellouin, N., Boucher, O., Sayer, A. M., Thomas, G. E., McComiskey, A., Feingold, G., Hoose, C., Kristjansson, J. E., Liu, X., Balkanski, Y., Donner, L. J., Ginoux, P. A., Stier, P., Grandey, B., Feichter, J., Sednev, I., Bauer, S. E., Koch, D., Grainger, R. G., Kirkevag, A., Iversen, T., Seland, O., Easter, R., Ghan, S. J., Rasch, P. J., Morrison, H., Lamarque, J. F., Iacono, M. J., Kinne, S., and Schulz, M.: Aerosol indirect effects - general circulation model intercomparison and evaluation with satellite data, Atmos. Chem. Phys., 9, 8697–8717, 2009.

Räisänen, P., Haapanala, P., Chung, C. E., Kahnert, M., Makkonen, R., Tonttila, J., and Nousiainen, T.: Impact of dust particle non-sphericity on climate simulations, Quarterly Journal of the Royal Meteorological Society, 139, 2222–2232, doi:10.1002/qj.2084, 2013.

Randles, C. A., Kinne, S., Myhre, G., Schulz, M., Stier, P., Fischer, J., Doppler, L., Highwood, E., Ryder, C., Harris, B., Huttunen, J., Ma, Y., Pinker, R. T., Mayer, B., Neubauer, D., Hitzenberger, R., Oreopoulos, L., Lee, D., Pitari, G., Di Genova, G., Quaas, J., Rose, F. G., Kato, S., Rumbold, S. T., Vardavas, I., Hatzianastassiou, N., Matsoukas, C., Yu, H., Zhang, F., Zhang, H., and Lu, P.: Intercomparison of shortwave radiative transfer schemes in global aerosol modeling: results from the AeroCom Radiative Transfer Experiment, Atmospheric Chemistry and Physics, 13, 2347–2379, doi:10.5194/acp-13-2347-2013, https://www.atmos-chem-phys.net/13/2347/2013/, 2013.

Regayre, L. A., Johnson, J. S., Yoshioka, M., Pringle, K. J., Sexton, D. M. H., Booth, B. B. B., Lee, L. A., Bellouin, N., and Carslaw, K. S.: Aerosol and physical atmosphere model parameters are both important sources of uncertainty in aerosol ERF, Atmospheric Chemistry and Physics, 18, 9975–10 006, doi:10.5194/acp-18-9975-2018, 2018.

Rossow, W. B. and Schiffer, R. A.: Advances in understanding clouds from ISCCP, Bull. Amer. Meteorol. Soc., 80, 2261–2288, doi:10.1175/1520-0477(1999)080<2261:AIUCFI>2.0.CO;2, 1999.

Salisbury, D. J., Anguelova, M. D., and Brooks, I. M.: On the variability of whitecap fraction using satellite-based observations, J. Geophys. Res. Oceans, pp. 6201—6222, doi:10.1002/2013JC008797, 2013.

Shindell, D. T., Lamarque, J.-F., Schulz, M., Flanner, M., Jiao, C., Chin, M., Young, P. J., Lee, Y. H., Rotstayn, L., Mahowald, N., Milly, G., Faluvegi, G., Balkanski, Y., Collins, W. J., Conley, A. J., Dalsoren, S., Easter, R., Ghan, S., Horowitz, L., Liu, X., Myhre, G., Nagashima, T., Naik, V., Rumbold, S. T., Skeie, R., Sudo, K., Szopa, S., Takemura, T., Voulgarakis, A., Yoon, J.-H., and Lo, F.: Radiative forcing in the

ACCMIP historical and future climate simulations, Atmospheric Chemistry and Physics, 13, 2939–2974, doi:10.5194/acp-13-2939-2013, 2013.

Smirnov, A., Holben, B. N., Slutsker, I., Giles, D. M., McClain, C. R., Eck, T. F., Sakerin, S. M., Macke, A., Croot, P., Zibordi, G., Quinn, P. K., Sciare, J., Kinne, S., Harvey, M., Smyth, T. J., Piketh, S., Zielinski, T., Proshutinsky, A., Goes, J. I., Nelson, N. B., Larouche, P., Radionov, V. F., Goloub, P., Krishna Moorthy, K., Matarrese, R., Robertson, E. J., and Jourdin, F.: Maritime Aerosol Network as a component of Aerosol Robotic Network, Journal of Geophysical Research: Atmospheres, 114, n/a–n/a, doi:10.1029/2008JD011257, d06204, 2009.

Sofiev, M., Soares, J., Prank, M., de Leeuw, G., and Kukkonen, J.: A regional-to-global model of emission and transport of sea salt particles in the atmosphere, Journal of Geophysical Research: Atmospheres, 116, doi:10.1029/2010JD014713, 2011.

Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., Salzmann, M., Schmidt, H., Bader, J., Block, K., Brokopf, R., Fast, I., Kinne, S., Kornblueh, L., Lohmann, U., Pincus, R., Reichler, T., and Roeckner, E.: Atmospheric component of the MPI-M Earth System Model: ECHAM6, Journal of Advances in Modeling Earth Systems, 5, 146–172, doi:10.1002/jame.20015, 2013.

Stevens, B., Fiedler, S., Kinne, S., Peters, K., Rast, S., Müsse, J., Smith, S. J., and Mauritsen, T.: MACv2-SP: a parameterization of anthropogenic aerosol optical properties and an associated Twomey effect for use in CMIP6, Geosci. Mod. Dev., 10, 433–452, doi:10.5194/gmd-10-433-2017, 2017.

Stier, P., Kinne, S., Kloster, S., Vignati, E., Wilson, J., Granzeveld, L., Teger, I., Werner, M., Balkanski, Y., Schulz, M., Boucher, O., Minikin, A., and Petzold, A.: The aerosol-climate model ECHAM5-HAM, Atmos. Chem. Phys., 5, 1125–1156, 2005.

Stier, P., Feichter, J., Roeckner, E., Kloster, S., and Esch, M.: The evolution of the global aerosol system in a transient climate simulation from 1860 to 2100, Atmospheric Chemistry and Physics, 6, 3059–3076, doi:10.5194/acp-6-3059-2006, 2006.

Stier, P., Seinfeld, J. H., Kinne, S., and Boucher, O.: Aerosol absorption and radiative forcing, Atmospheric Chemistry and Physics, 7, 5237–5261, doi:10.5194/acp-7-5237-2007, 2007.

Stier, P., Schutgens, N. A. J., Bellouin, N., Bian, H., Boucher, O., Chin, M., Ghan, S., Huneeus, N., Kinne, S., Lin, G., Ma, X., Myhre, G., Penner, J. E., Randles, C. A., Samset, B., Schulz, M., Takemura, T., Yu, F., Yu, H., and Zhou, C.: Host model uncertainties in aerosol radiative forcing estimates: results from the AeroCom Prescribed intercomparison study, Atmospheric Chemistry and Physics, 13, 3245–3270, doi:10.5194/acp-13-3245-2013, 2013.

Struthers, H., Ekman, A. M. L., Glantz, P., Iversen, T., K irkevåg, A., Mårtensson, E. M., Seland, Ø., and Nilsson, E. D.: The effect of sea ice loss on sea salt aerosol concentrations and the radiative balance in the Arctic, Atmospheric Chemistry and Physics, 11, 3459–3477, doi:10.5194/acp-11-3459-2011, https://www.atmos-chem-phys.net/11/3459/2011/, 2011.

Tegen, I., Harrison, S., Kohfeld, K., Prentice, I., Coe, M., and Heimann, M.: Impact of vegetation and preferential source areas on global dust aerosols: Results from a model study, J. Geophys. Res., 107, 4576, 2002.

Tegen, I., Neubauer, D., Ferrachat, S., Siegenthaler-Le Drian, C., Bey, I., Schutgens, N., Stier, P., Watson-Parris, D., Stanelle, T., Schmidt, H., Rast, S., Kokkola, H., Schultz, M., Schroeder, S., Daskalakis, N., Barthel, S., Heinold, B., and Lohmann, U.: The aerosol-climate model ECHAM6.3-HAM2.3: Aerosol evaluation, Geoscientific Model Development Discussions, 2018, 1–54, doi:10.5194/gmd-2018-235, 2018.

van Noije, T. P. C., Le Sager, P., Segers, A. J., van Velthoven, P. F. J., Krol, M. C., Hazeleger, W., Williams, A. G., and Chambers, S. D.: Simulation of tropospheric chemistry and aerosols with the climate model EC-Earth, Geosci. Model Dev., 7, 2435–2475, doi:10.5194/gmd-7-2435-2014, 2014.

Walters, D., Boutle, I., Brooks, M., Melvin, T., Stratton, R., Vosper, S., Wells, H., Williams, K., Wood, N., Allen, T., Bushell, A., Copsey, D., Earnshaw, P., Edwards, J., Gross, M., Hardiman, S., Harris, C., Heming, J., Klingaman, N., Levine, R., Manners, J., Martin, G., Milton, S.,

Mittermaier, M., Morcrette, C., Riddick, T., Roberts, M., Sanchez, C., Selwood, P., Stirling, A., Smith, C., Suri, D., Tennant, W., Vidale, P. L., Wilkinson, J., Willett, M., Woolnough, S., and Xavier, P.: The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations, Geoscientific Model Development, 10, 1487–1520, doi:10.5194/gmd-10-1487-2017, 2017.

Wilson, D. R., Bushell, A. C., Kerr-Munslow, A. M., Price, J. D., and Morcrette, C. J.: PC2: A prognostic cloud fraction and condensation scheme. I: Scheme description, Q. J. R. Meteorol. Soc, p. 2093–2107, doi:https://doi.org/10.1002/qj.333, 2008.

Zhang, K., O'Donnell, D., Kazil, J., Stier, P., Kinne, S., Lohmann, U., Ferrachat, S., Croft, B., Quaas, J., Wan, H., Rast, S., and Feichter, J.: The global aerosol-climate model ECHAM-HAM, version 2: sensitivity to improvements in process representations, Atmospheric Chemistry and Physics, 12, 8911–8949, doi:10.5194/acp-12-8911-2012, 2012.

Zhang, K., Wan, H., Liu, X., Ghan, S. J., Kooperman, G. J., Ma, P.-L., Rasch, P. J., Neubauer, D., and Lohmann, U.: Technical Note: On the use of nudging for aerosol-climate model intercomparison studies, Atmospheric Chemistry and Physics, 14, 8631–8645, doi:10.5194/acp-14-8631-2014, 2014.

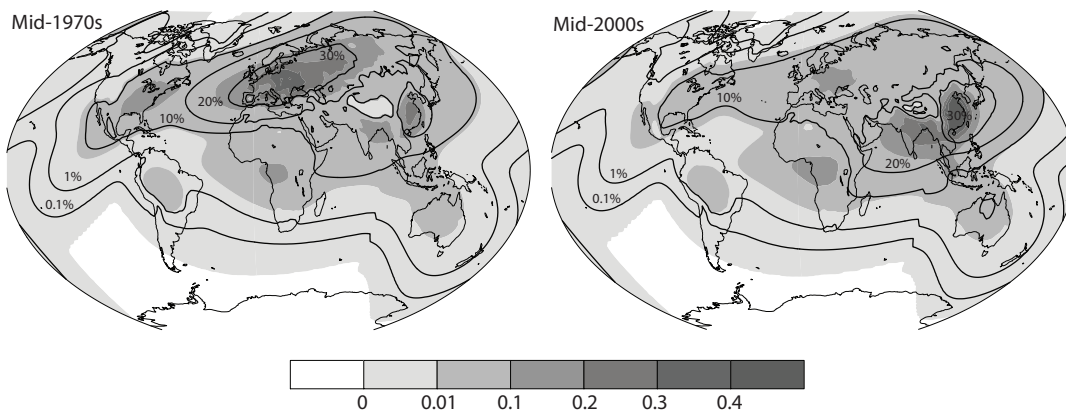22

Figure 1. Mean anthropogenic aerosol optical depth (τ_a, shaded) and fractional increase in cloud droplet number (η_N, contours) associated with anthropogenic aerosol. Shown are annual means of τ_a at 550 nm and η_N for the (left) mid-1970s and (right) mid-2000s from MACv2-SP that prescribes annually repeating monthly maps of τ_a in the participating models. Note the non-linear scale.
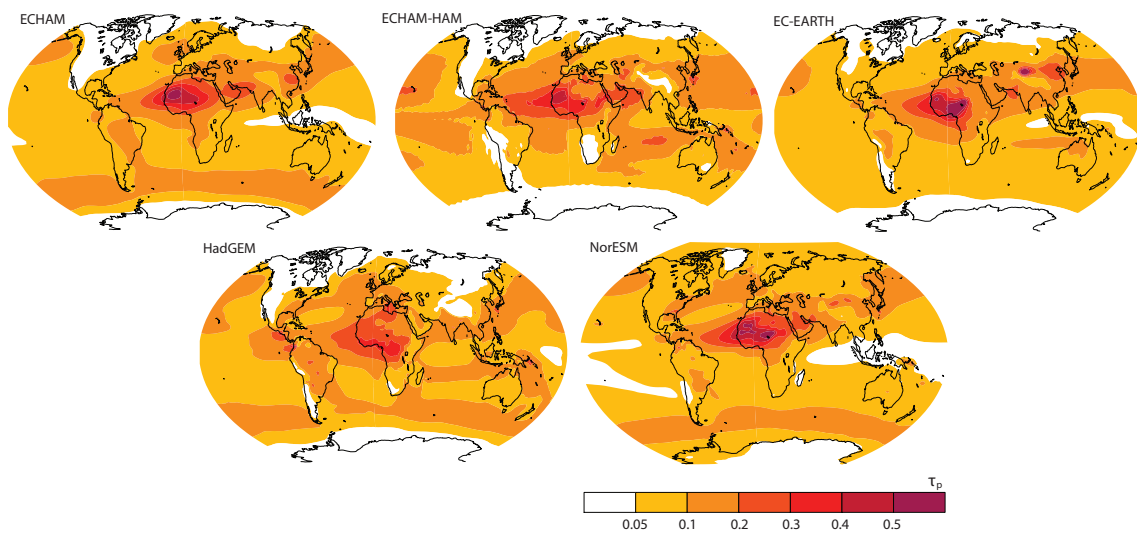
Figure 2. Mean pre-industrial aerosol optical depth (τ_p). Shown are annual means of τ_p of the radiation band around 550 nm for each model.
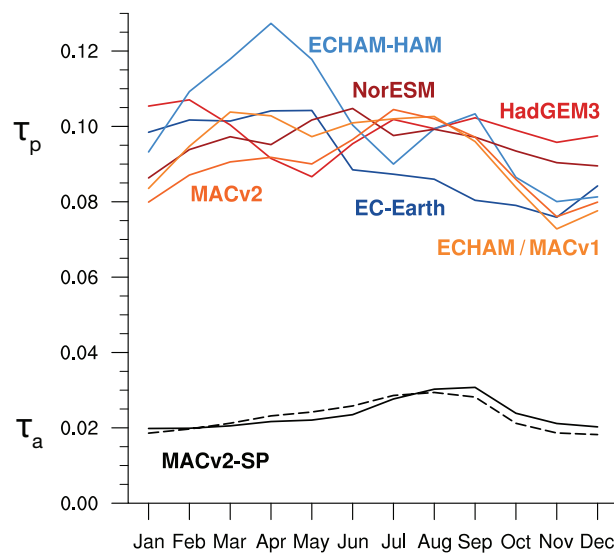
Figure 3. Annual cycle of the global mean aerosol optical depth at 550 nm. Shown are monthly means of (colors) τ_p from the models and (black) τ_a for the (dashed) mid-1970s and (solid) mid-2000s from MACv2-SP.
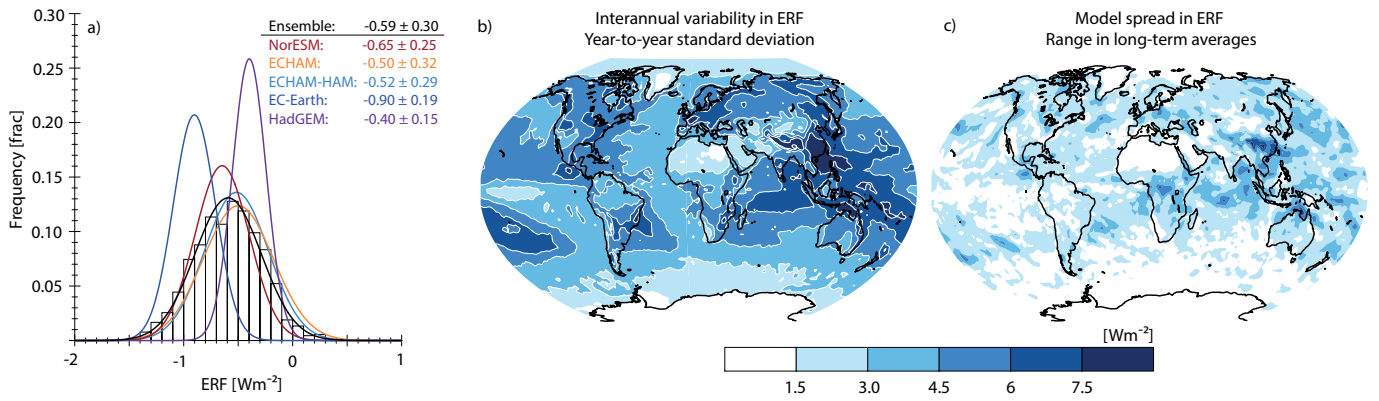
Figure 4. Variability in annual ERF estimates for the mid-2000s. Panel (a) shows Gaussian distributions of annual ERF estimates for present-day from (colors) individual model ensembles and (black) the entire multi-model, multi-member ensemble. The bars are the frequency histogram of one-year ERF estimates from all models, and the legend indicates the means and standard deviations of the ERF estimates. Panel (b) shows the regional standard deviation of annual contributions to ERF from the entire multi-model, multi-member ensemble as measure for the interannual variability internal to the model ensemble. Panel (c) shows the range in the long-term averaged ERFs of the models as measure for the spread in ERF associated with model differences. ERF are for the shortwave (SW) spectrum at the top of atmosphere (TOA) for all-sky conditions.
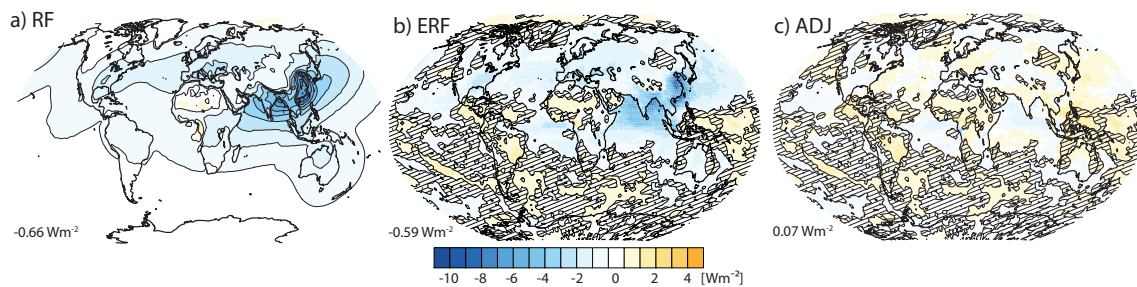
Figure 5. Multi-model, multi-member ensemble mean of the anthropogenic aerosol radiative effects for the mid-2000s. Shown are the (a) instantaneous and (b) effective radiative forcing as well as (c) the net contribution from rapid adjustments for SW at the TOA in all-sky conditions. Hatching in (b, c) indicates non-significant values at a 10% significance level. The numbers in the lower left corner are the spatial averages. The ensemble-mean RF is averaged over three climate models, the ensemble-mean ERF over five climate models, and the ensemble-mean adjustment is their difference.
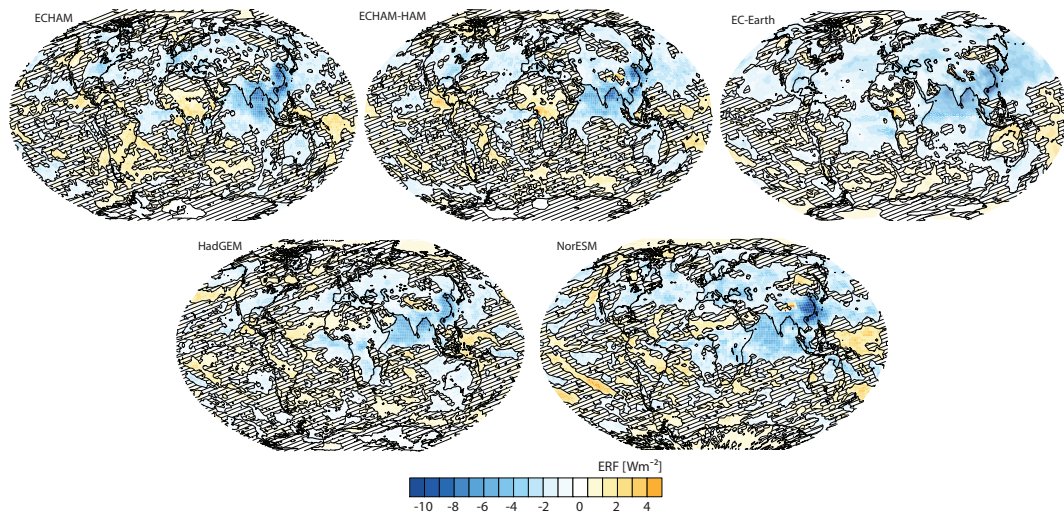
Figure 6. Multi-member ensemble mean of effective radiative effects of anthropogenic aerosol for the mid-2000s. Shown are the effective radiative forcing for SW at the TOA in all-sky conditions for each model. Hatching indicates non-significant values at a 10% significance level.
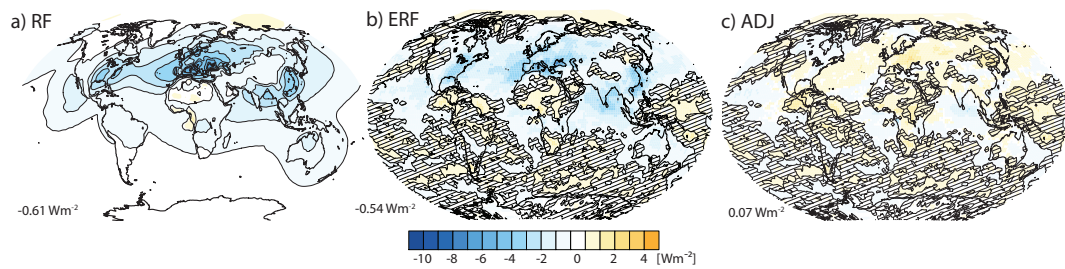
a) RF

-0.61 Wm⁻²

b) ERF

-0.54 Wm⁻²

-10 -8 -6 -4 -2 0 2 4 [Wm⁻²]

c) ADJ

0.07 Wm⁻²

Figure 7. Multi-model, multi-member ensemble mean of the anthropogenic aerosol radiative effects for the mid-1970s. As Figure 5, but with the anthropogenic aerosol pattern of the mid-1970s.
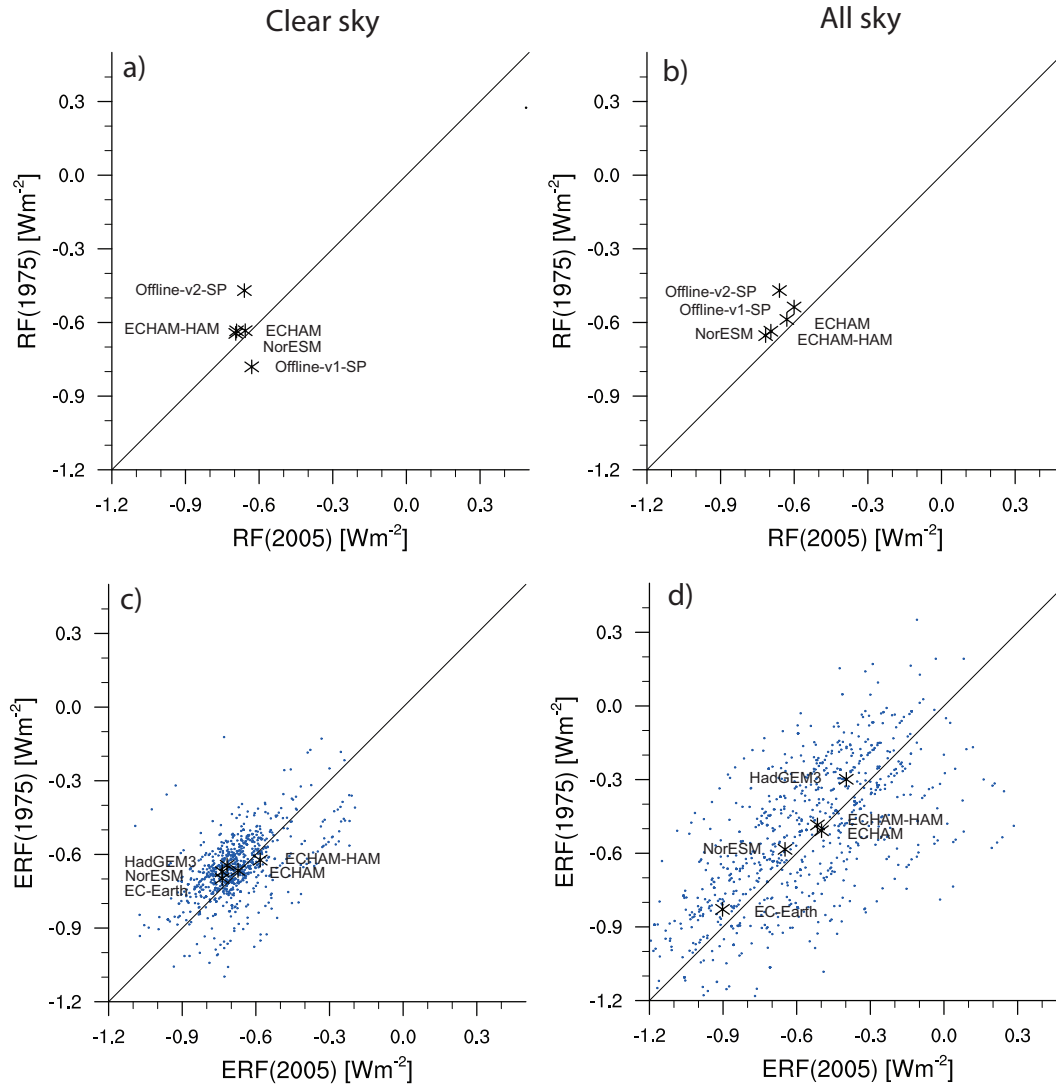
Figure 8. Anthropogenic aerosol forcing of the mid-1970s against the mid-2000s. Shown are the (top) instantaneous and (bottom) effective radiative forcing for SW at the TOA from the pollution of the mid-1970s against the mid-2000s for (left) clear and (right) all sky. Thick crosses are the ensemble means. Blue dots in (c, d) are the model averages of individual years representing the year-to-year variability internal to the model ensemble.
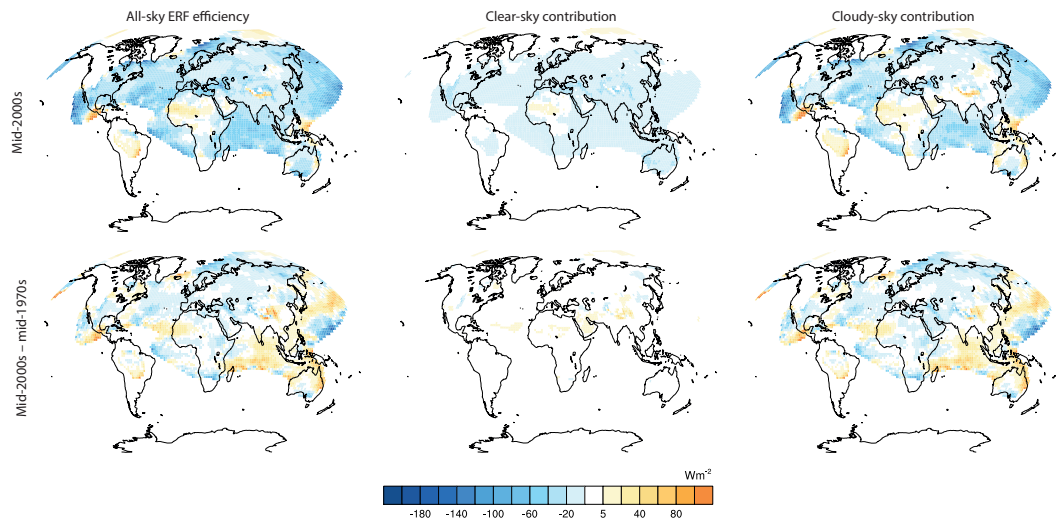
Figure 9. Anthropogenic aerosol effective radiative forcing efficiencies, in W m^{-2} per unit optical depth, for (left) all-sky, (middle) clear-sky, and (right) cloudy-sky. The top row shows efficiencies for mid-2000s anthropogenic aerosols. The bottom row shows differences made by using the pattern for the mid-1970s.
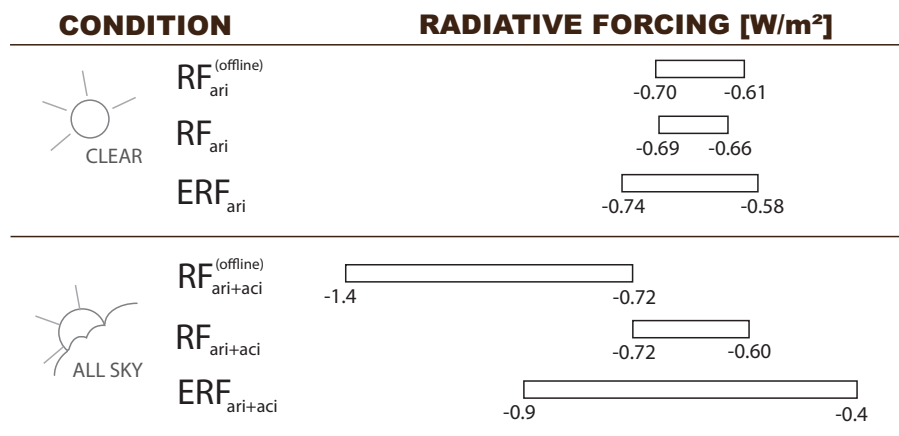
Figure 10. Summary of model spread in anthropogenic aerosol forcing for the mid-2000s. Shown are the instantaneous (RF) and effective radiative forcing (ERF) of aerosol-radiation and aerosol-cloud interactions for the shortwave spectrum at the top of the atmosphere for clear and all sky from Tab.2. The RF from the offline radiation-transfer calculations consider additional uncertainty sources and are shown as separate bars. Refer to Section 2.1 for details.
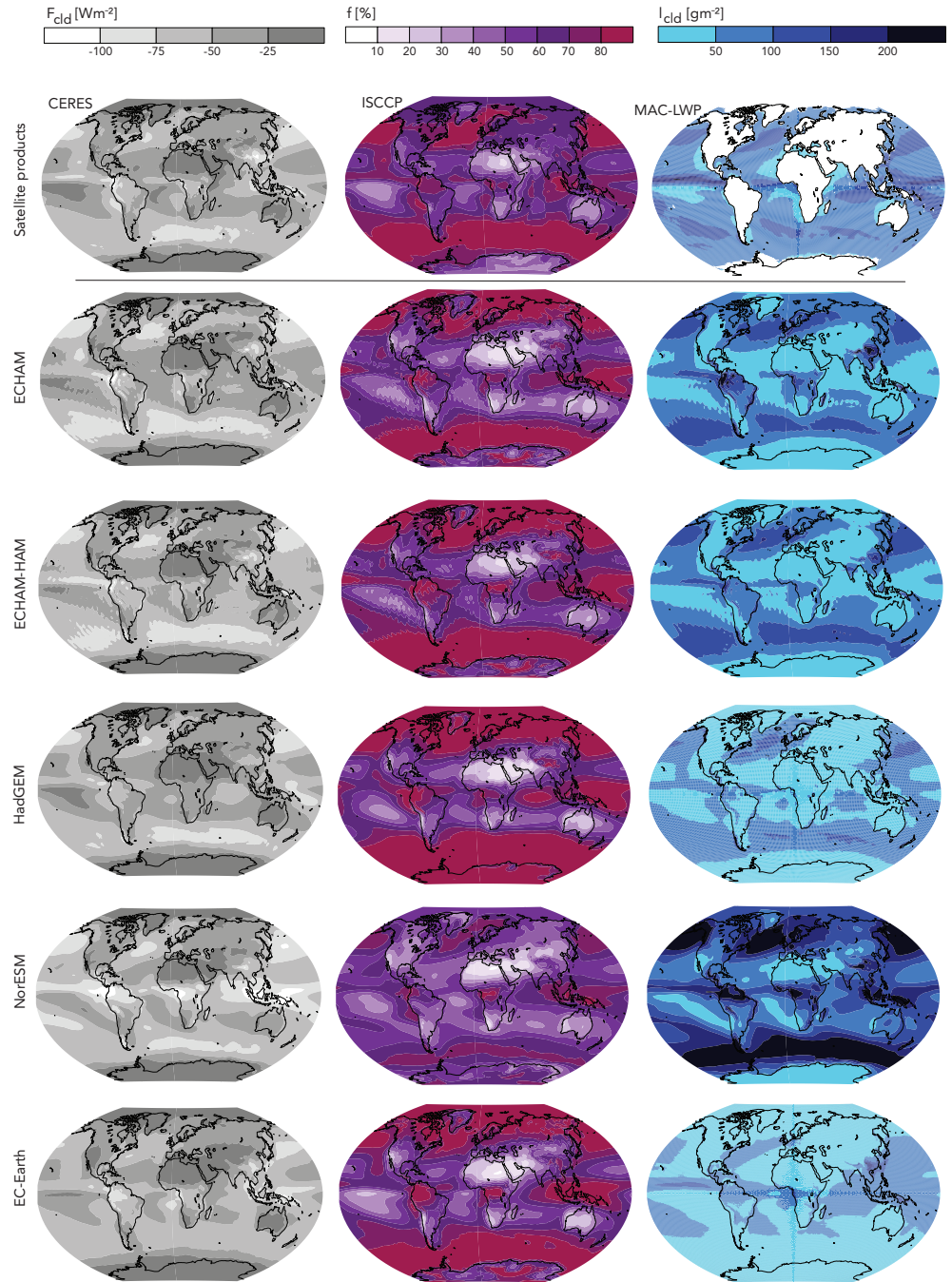
Figure A1. Multi-member ensemble means of cloud characteristics for the mid-2000s compared to climatologies derived from satellite observations (Table A1). Shown are the mean (left column) SW cloud radiative effect at the TOA, F_{cld}, (middle column) total cloud cover, f, and (right column) liquid water path, l_{cld} from (top row) the satellite products and (rows beneath) the models. Areas without available ~~satellite~~ data are shaded white.
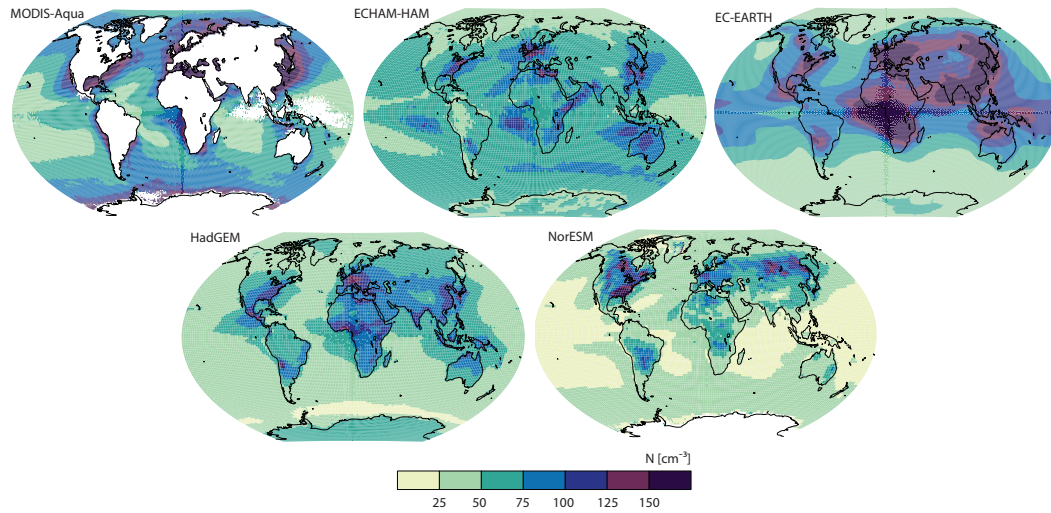
Figure A2. In-cloud droplet number concentration for the mid-2000s. Shown are the annually and vertically averaged in-cloud droplet number concentration (N) from the aerosol-climate models and from the MODIS satellite product by Bennartz and Rausch (2017). Areas without available ~~satellite~~ data are shaded white.
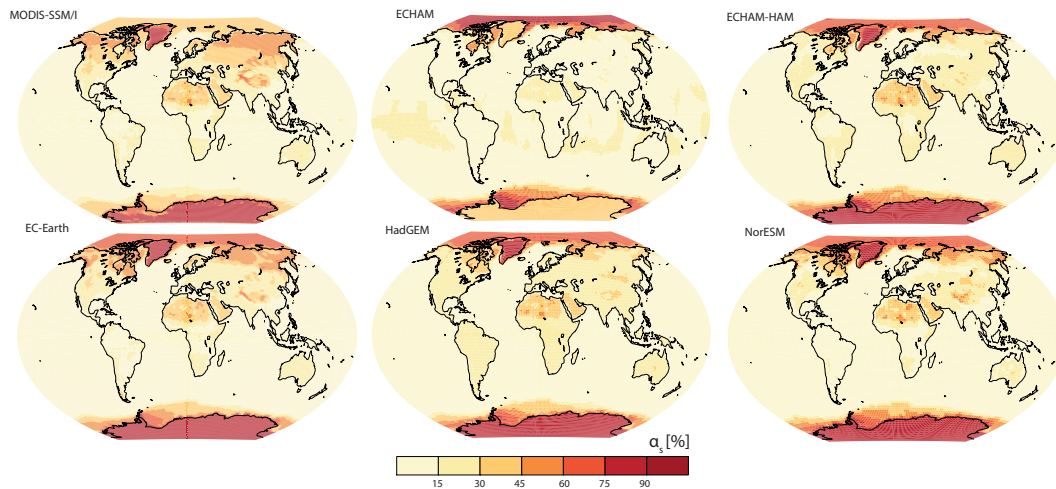

Figure A3. Surface albedo for shortwave radiation for the mid-2000s. Shown are the mean surface albedo for shortwave radiation (α_s) from the models and the satellite product from Kinne et al. (2013).

Table 1. Model experimental setup

Model	Horizontal resolution (longitude x latitude)	Number of vertical levels	Pre-industrial aerosol (1850)	Anthropogenic aerosol (increase since 1850)
ECHAM	1.875° x 1.875°	47	MACv1 climatology	MACv2-SP
ECHAM-HAM	1.875° x 1.875°	47	Online	MACv2-SP
EC-Earth	1.875° x 1.875°	91	TM5 climatology	MACv2-SP
HadGEM3	1.875° x 1.25°	85	HadGEM3 climatology	MACv2-SP
NorESM	2.5° x 1.894°	26	Online	MACv2-SP
Offline-v2-SP	1° x 1°	20	MACv2	MACv2-SP
Offline-v1-SP	1° x 1°	20	MACv1	MACv2-SP
Offline-v2	1° x 1°	20	MACv2	MACv2
Offline-v1	1° x 1°	20	MACv1	MACv1

35

Table 2. Ensemble averages of the shortwave instantaneous (RF) and effective (ERF) radiative forcing, and net contribution from rapid adjustments (ADJ) at the surface (SFC) and the top of the atmosphere (TOA) for all sky (clear sky) in Wm^{-2} for the period 1850 to 2005. The first block shows aerosol-climate models with MACv2-SP, and the second block shows estimates of the offline radiative transfer model.

	RF_{SFC}	RF_{TOA}	ERF_{TOA}	ADJ_{TOA}
ECHAM	-1.52 (-1.64)	-0.60 (-0.66)	-0.50 (-0.67)	0.1 (-0.01)
ECHAM-HAM	-1.63 (-1.67)	-0.72 (-0.69)	-0.52 (-0.58)	0.2 (0.11)
EC-Earth	/	/	-0.90 (-0.74)	/
HadGEM3	/	/	-0.40 (-0.72)	/
NorESM	-1.46 (-1.60)	-0.68 (-0.68)	-0.65 (-0.74)	0.03 (-0.06)
Offline-v2-SP	-1.8 (-1.7)	-0.75 (-0.62)	/	/
Offline-v1-SP	-1.7 (-1.6)	-0.72 (-0.61)	/	/
Offline-v2	-2.3 (-1.9)	-1.1 (-0.70)	/	/
Offline-v1	-2.7 (-2.0)	-1.4 (-0.63)	/	/

Table A1. Gridded climatologies of satellite retrievals used for model evaluation.

Name	Description	Variable	Time
CERES	Energy balanced and filled data of the Clouds and the Earth's Radiant Energy System, Ed. 4 (Loeb et al., 2009)	Cloud shortwave radiative effects at the top of the atmosphere, F_{cld} [Wm^{-2}]	$2001-2014$
ISCCP	International Satellite Cloud Climatology Project (Rossow and Schiffer, 1999)	Total cloud cover, f [%]	$1983-2009$
MAC-LWP	Multi-sensor Advanced Climatology (Elsaesser et al., 2017) (Elsaesser et al., 2016, 2017)	Liquid water path, l_{cld} [gm^{-2}]	$2000-2016$
MODIS	Climatology based on Moderate Resolution Imaging Spectroradiometer aboard Aqua (Bennartz and Rausch, 2017)	Cloud droplet number concentration in warm clouds, N [cm^{-3}]	$2003-2015$
MODIS-SSM/I	Climatology based on Moderate Resolution Imaging Spectroradiometer and microwave data (Kinne et al., 2013)	Surface albedo for shortwave radiation, α_s [%]	$1987-2007$

Table A2. Global mean statistics for clouds, aerosols and surface albedo. The numbers given for l_{cld} and N are averages over ocean regions, consistent with the satellite data availability (Figures A1 and A2). Details on the satellite products are listed in Table A1.

	F_{cld} [Wm^{-2}]	f [%]	l_{cld} [gm^{-2}]	N [cm^{-3}]	τ_p	α_s [%]
ECHAM	-47.5	63	65	84	0.093	16
ECHAM-HAM	-49.1	68	69	65	0.097	15
EC-Earth	-46.2	65	42	91	0.091	15
HadGEM3	-44.3	69	57	56	0.098	15
NorESM	-55.5	55	133	34	0.096	14
Satellite retrieval	-45.8	66	82	77	-	15