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], Ken Carslaw[7], Risto Makkonen[3,8], 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]School of Earth and Environment, University of Leeds, Leeds, UK
[8]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. The radiative forcing of anthropogenic aerosol remains a key uncertainty in the understanding of climate change. This study quantifies the model spread in aerosol forcing associated with (i) variability internal to the atmosphere and (ii) differences in the model representation of weather. We do so by performing ensembles of atmosphere-only simulations with five state-of-the-art Earth system models, four of which will be used in the sixth coupled model inter-comparison project (CMIP6, Eyring et al., 2016). We calculate the instantaneous radiative forcing (RF), effective radiative forcing (ERF), and rapid adjustments by comparing 10-year long ensemble simulations with aerosol distributions for 1850, the mid-1970s and the mid-2000s. The complexity of the anthropogenic aerosol is herein reduced by prescribing the same annually-repeating monthly patterns of the anthropogenic aerosol optical properties and associated effects on the cloud droplet number concentration. We quantify a relatively small model spread in the long-term averaged ERF compared to the overall possible range in annual ERF estimates associated with variability internal to the model ensemble. This implies that identifying the true model spread in ERF associated with differences in the representation of meteorological processes and natural aerosol requires averaging over a sufficiently large number of annual estimates. Despite major inter-model differences in natural aerosol and clouds, all models show only a small change in the global-mean ERF due to the substantial change in the global anthropogenic aerosol distribution between the mid-1970s and mid-2000s, the ensemble mean ERF being -0.54 Wm^{-2} for the mid-1970s and -0.59 Wm^{-2} for the mid-2000s. This result suggests that inter-comparing ERF changes between two periods rather than absolute magnitudes relative to pre-industrial might provide a better test for a model's ability for representing climate evolutions.

1 Introduction

Despite decades of research on the radiative forcing of anthropogenic aerosol, quantifying the present-day magnitude and reconstructing the historical evolution of the forcing remains challenging. Recent work has indicated that natural variability affects estimates of the effective radiative forcing (ERF) of anthropogenic aerosol (Fiedler et al., 2017). More specifically, natural variability was identified as a cause for increases and decreases in the global mean ERF associated with the spatial change in anthropogenic AOD (τ_a) between the mid-1970s and mid-2000s. The anthropogenic aerosol pollution in the mid-1970s was herein 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 changes in the global ERF over time.

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), a single model as used in Fiedler et al. (2017) 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, reflected by the different spatial pattern of τ_a shown in Fig. 1, affect the global magnitude of ERF?" using ensembles of simulations from five global aerosol-climate models with reduced aerosol complexity. In this context, we additionally ask: "What is the relative contribution of variability amongst and within models to the spread in ERF?", and document the model diversity for the pre-industrial aerosol and cloud characteristics that are relevant for ERF of anthropogenic aerosol. Such model differences have previously been identified for other climate models (e.g., Nam et al., 2012; Fiedler et al., 2016; Crueger et al., 2018).

We address our research questions by reducing the complexity of the anthropogenic aerosol representation in an ensemble of modern aerosol-climate models. Previously a reduction of 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 (Fiedler et al., 2017; Stevens et al., 2017), but keep the full model diversity in other aspects. It allows us to eliminate the uncertainties in process modelling of anthropogenic aerosol and focus on the uncertainties in other processes influencing the radiative forcing. In other words, prescribing identical anthropogenic aerosol optical properties and an associated effect on the cloud droplet number concentration across models allows us to study those sources of uncertainty that remain if we pretend to know the spatial distribution of anthropogenic aerosol. We can thereby quantify the sole impact of other model differences, such as the natural aerosol, meteorology, radiative transfer, and surface albedo, on the radiative forcing of observationally constrained anthropogenic aerosol in a state-of-the-art multi-model context. As such our model inter-comparison with MACv2-SP 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).

Throughout our model inter-comparison, we consider the effect of model-internal variability on the magnitude of ERF by producing equally-sized ensembles of simulations for all participating models. Model-internal variability is herein measured as the year-to-year variations internal to the model that are associated with the changing weather. We provide observational

benchmarks for the inter-comparison of the complex models with satellite data and results from a stand-alone atmospheric radiation transfer model for quantifying differences in the instantaneous radiative forcing (RF). 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 at the end of the article.

2   Method

2.1   Participating Models

This work uses five Earth-system models and one stand-alone radiation transfer code. The participating models are the atmosphere component ECHAM6.3 of the Earth system model MPI-ESM1.2 of the Max-Planck Institute for Meteorology (MPI-M, Mauritsen and et al., in review), ECHAM6.3-HAM2.3 from the ETH Zürich (Tegen et al., 2018; Neubauer et al., submitted), EC-Earth (e.g., Hazeleger et al., 2010; Döscher and et al., in prep.) run here at the Royal Netherlands Meteorological Institute, NorESM (Bentsen et al., 2013; Iversen et al., 2013; Kirkevåg et al., 2013) run for the present study at the Finnish Meteorological Institute, and HadGEM3 (Walters et al., 2017) developed at the UK Met Office. All models except MPI-ESM1.2 can treat aerosol and their interaction with meteorological processes with complex bottom-up parameterisation schemes linking aerosols to radiation and clouds. The model differences in radiation, cloud, and aerosol physics packages that are relevant to anthropogenic aerosol forcing are summarised in Appendix A.

In the present study, we prescribe identical optical properties of anthropogenic aerosols for representing aerosol-radiation interactions (F_{ari}) and an associated change in the cloud droplet number concentration (N) for representing aerosol-cloud interactions (F_{aci}) by implementing MACv2-SP (Fiedler et al., 2017; Stevens et al., 2017) in all models. All other aspects remain model-dependent, e.g., the treatment of the pre-industrial aerosol and clouds (Appendix A). MACv2-SP mimics the spatio-temporal distribution and wavelength dependence of anthropogenic aerosols as well as a change in the cloud droplet number concentration to induce aerosol radiative effects 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.). The single scattering albedo is 0.93 for industrial plumes and 0.87 for plumes with seasonally active biomass burning. The asymmetry parameter is set to 0.63. Here, we use MACv2-SP with the CMIP6 reconstructed evolution of anthropogenic aerosol emission, identical to the one used by Fiedler et al. (2017).

Figure 1 shows the annual mean patterns of the prescribed anthropogenic aerosol optical depth (τ_a), and the percentage increase in the cloud droplet number concentration (η_N) relative to the pre-industrial level for the mid-1970s and mid-2000s. EC-Earth accounts for F_{aci} by multiplying the pre-industrial N in the cloud microphysics with η_N from MACv2-SP. All other models represent F_{aci} in the form of a Twomey effect by multiplying η_N with N prior to the radiation transfer calculation. The effective parameter η_N herein increases the cloud reflectivity of the shortwave radiation.

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). Regional differences occur primarily over oceans and deserts, where observations are typically sparse. It is herein 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) compared to the monthly mean climatology MACv1 in ECHAM. 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 bottom-up aerosol modelling (e.g., Räisänen et al., 2013; Evan et al., 2014; Huneeus et al., 2016).

In addition to the complex Earth system models, we use the offline radiation code of Kinne et al. (2013) with eight solar and twelve infrared bands for an observational benchmark of the instantaneous radiative forcing. The code reads monthly maps of the atmospheric and surface properties. These are monthly means for the cloud properties from ISCCP, surface albedo from MODIS, and surface temperature from AeroCom, described in detail by Kinne et al. (2013). The radiative transfer calculation considers nine different sun elevations and eight random permutations of cloud heights and overlap. Aerosol column properties at 550 nm are defined by the MPI-M's Aerosol Climatology (MAC). We calculate the radiation transfer with both MAC version one (MACv1, Kinne et al., 2013) and two (MACv2, Kinne, submitted), the latter of which considers more recent observational data, e.g., from the Maritime Aerosol Network (MAN, Smirnov et al., 2009), and a different temporal evolution of the anthropogenic aerosol fraction. MACv1 produces a temporal scaling of the anthropogenic aerosol fraction based on the emission inventory by Dentener et al. (2006), while MACv2 uses the one by Lamarque et al. (2010). The two climatologies differ in their pre-industrial aerosol burden, namely a lower background burden representative for 1750 is used in MACv1 in contrast to the background for 1850 in MACv2. The mean annual cycle of the pre-industrial aerosol optical depths of MACv1 and MACv2 is shown in Figure 3, along with the pre-industrial aerosol optical depths from the other participating models. The aerosol vertical distribution and fine-mode anthropogenic fraction of AOD are derived from global models participating in AeroCom (e.g., Myhre et al., 2013).

2.2 Experiment strategy

All experiments are carried out with the atmosphere-only model configurations with 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 other model components than MACv2-SP, and choosing both the boundary and initialization data like they usually do. 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, each participating model was free to individually set up all other aspects than the anthropogenic aerosol treatment. We therefore keep for instance the model diversity for the physical parameterisations of radiation and clouds (Appendix A).

We produce ensembles of simulations from each model motivated by the effect of natural variability on ERF estimates in ECHAM (Fiedler et al., 2017). For each model, we perform a total of 12 experiments with prescribed sea ice and sea-surface temperature for the years 2000−2010 inclusive. These are six experiments with pre-industrial aerosol optical depth (τ_p) as of

the year 1850, three experiments with τ_p and anthropogenic aerosol from MACv2-SP for the year 1975, and three experiment with τ_p and anthropogenic aerosol from MACv2-SP for the year 2005. The first year of each 11-year run is considered as a spin-up period and is excluded from the analysis, thus all analyses are for the period $2001-2010$. We have chosen the ten-year period for including variability in the boundary conditions.

5   The instantaneous radiative forcing (RF) of anthropogenic aerosol in clear and all sky is estimated from double radiation calls in the models having this functionality, i.e., by calculating the atmospheric transfer of shortwave radiation once with and once without the anthropogenic aerosol optical properties and an associated effect on the cloud droplet number concentration from MACv2-SP. The reference aerosol was herein for the year 1850. This gives us in total 30 annual estimates of RF per model for each of the two pollution patterns (Figure 1), which is sufficient for a precise estimate of RF and can be directly

10   compared to the offline radiative transfer calculations.

The effective radiative forcing (ERF) is calculated relative to pre-industrial simulations for each model by subtracting the monthly mean shortwave radiation budgets and producing annual averages. This approach is chosen for illustrating the effect of year-to-year variability on ERF estimates. Since we are using MACv2-SP, the ERF estimates account for F_{ari} and F_{aci}. The latter includes the Twomey effect and the radiative effect of rapid adjustments in clouds, atmosphere and surface properties.

15   Subtracting the time series of the six pre-industrial experiments from each of the three experiments with additional anthropogenic aerosol adds up to 6x3 time series of monthly ERF estimates over ten years per model, i.e., 180 annual estimates per model and τ_a pattern in total. We choose annually averaged ERF for estimating the impact of natural variability internal to the atmosphere for each model. The long-term averaged ERFs over 180 years are used for identifying systematic model differences in ERF. Such long time periods are sufficient for diagnosing ERF, e.g., in the here participating model ECHAM (Fiedler et al.,

20   2017), and the RFMIP protocol recommends a thirty-year average for diagnosing the ERF of a model (Pincus et al., 2016). Additionally, we calculate the net contribution of rapid adjustments (ADJ) to ERF by subtracting RF from ERF for each model.

3 Results

3.1 Spread in present-day ERF

We characterise the spread in the effective radiative forcing (ERF) from the model ensemble, summarised in Table 2. For

25   doing so, we first calculate the multi-model mean as a reference value. The all-sky ERF at the TOA for the entire multi-model, multi-member ensemble is -0.59 Wm^{-2} with a year-to-year standard deviation of 0.3 Wm^{-2} translating to a typical percentage variability of roughly 50%. The entire range in annual ERFs including model-internal variability is -1.5 Wm^{-2} to +0.5 Wm^{-2}. The cloud masking effect, i.e., here going from clear to all-sky conditions (Table 2), reduces ERF at TOA by $10-50\%$ in all models except EC-Earth and is most pronounced in HadGEM3. In EC-Earth, the stronger F_{aci}, due to both the Twomey

30   and cloud-lifetime effects, overcompensates the cloud-masking effect, i.e., the all-sky ERF is more negative than the value for clear-sky conditions.

The long-term mean estimates of ERF are similar in ECHAM and ECHAM-HAM, despite the model differences in representing τ_p (Section 2.1). This suggests that the use of a prescribed climatology of τ_p (ECHAM) and an interactive simulation

of τ_p (ECHAM-HAM) are similarly useful for determining the model-mean ERF of anthropogenic aerosol. More generally, this hints that the ERF of anthropogenic aerosol is not strongly sensitive to the variability of natural aerosol not prescribed by the monthly climatology, when the aerosol-cloud interaction is treated consistently.

The year-to-year variability in ERF is illustrated by the Gaussian distribution fitted to the frequency histogram in Fig. 4a. Compared to the internal variability of the entire multi-model ensemble, the multi-model spread in the ensemble mean ERF of individual models is smaller, with a range of -0.40 Wm^{-2} to -0.9 Wm^{-2}. This multi-model spread corresponds to a range in differences to the multi-model mean of just -0.31 Wm^{-2} to +0.19 Wm^{-2}. This spread is even smaller when we exclude the ERF of EC-Earth that also represents cloud-lifetime effects, additional to the Twomey effect simulated in all models. The rather small multi-model spread is astonishing since although the models treat the anthropogenic aerosol and Twomey effect consistently, they differ in all other aspects, including the representation of clouds that is documented in the Appendix B.

What does the large variability imply for model-based estimates of ERF? If one wants to quantify model differences in ERF, it is essential to base the estimate for each model on a sufficiently large number of simulated years, i.e., either with sufficiently (i) long simulations with annually repeating aerosol or (ii) many simulations with transient changes. Otherwise one could not determine whether the ERF estimates are representative for the long-term averaged values. Given the similar year-to-year variability in ERF in the models, the precision of ERF estimates from ECHAM (Fiedler et al., 2017) is a reasonable approximation for the whole ensemble of models in this study.

3.2 Regional contributions to ERF

Regional contributions to ERF for the mid-2000s are shown as ensemble averages and for each model in Figure 5 and 6, respectively. The largest contributions to ERF are found over East Asia, consistent with the regional maximum in τ_a (Figure 5b). The general time-mean pattern of radiative effects is similar in the models. Distinct regions, however, show differences in the magnitude and detectability of the contributions to ERF, e.g., in central Africa where the radiative effects range from positive to negative. Consequently, the ensemble averaged contribution to ERF for that region is small.

Another interesting example for differences in regional contributions to ERF is the North Atlantic where efforts are made to use a volcanic eruption in Iceland to constrain radiative effects of anthropogenic aerosol (Malavelle et al., 2017). 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 away from Iceland (Figure 5). Close to Iceland, the ERF is generally close to the limit of detectability.

These regional model differences in ERF paired with year-to-year variability (Figure 4) suggest that more than one model ensemble would be needed for constraining the radiative effect of anthropogenic aerosol. Irrespectively whether we compute the standard deviation for the all-sky ERF for the aerosol pattern of the mid-1970s or the mid-2000s, the pattern and strength of the regional natural variability in ERF is robust (not shown). This implies that even for a larger perturbation of the tropospheric aerosol burden like in the mid-1970s over the North Atlantic, the natural variability of the atmosphere is a hurdle in constraining the regional radiative effect in addition to model differences in radiative effects.

The regional model spread in contributions to ERF are typically smaller than the differences associated with natural variability in the model ensemble (Figure 4b−c). However, the models disagree on the exact magnitude of the forcing in some regions, e.g., in the relatively large anthropogenic perturbation in East Asia for the mid-2000s (Figure 4c). The natural variability paired with the systematic model differences in the radiative effects suggests that a multi-model and multi-simulation ensemble is necessary for determining a regional climatological mean radiative effect of anthropogenic aerosol.

3.3 Benchmarking RF

We use offline radiation transfer calculations for providing benchmarks for the instantaneous radiative forcing (RF) of the complex models (Section 2.1). As a first step, we decompose ERF of the complex models into RF and the net contribution of rapid adjustments. RF is determined through double calls to the radiation calculation in each model and is considerably less variable than ERF (Figure 5). The net contribution of rapid adjustments to the global mean ERF ranges from -0.6 Wm^{-2} (EC-Earth) to 0.2 Wm^{-2} (ECHAM-HAM) at TOA, and acts to weaken the forcing magnitude in most models. RF clearly dominates the ERF magnitude in all models that use η_N in the radiation transfer calculation (Table 2). It is worthwhile recalling that these models consider F_{aci} in the form of a Twomey effect. When additionally cloud-lifetime effects are represented with MACv2-SP (EC-Earth), the net contribution from rapid adjustments can become significantly larger.

We compare the climate-model estimates of RF with offline radiation transfer calculations that use satellite observations of the atmosphere and surface following the method of Kinne et al. (2013), and the MACv2-SP aerosol (Section 2.1). The offline estimated all-sky RF with MACv2-SP (Offline-v1-SP and Offline-v2-SP) are in close agreement with the complex models that represent F_{aci} in form of a Twomey effect. This agreement is remarkable since the aerosol-climate models and the offline model differ in many aspects, including the representation of clouds that is documented in the Appendix B. The more negative clear-sky RF at the TOA in all complex models compared to the offline estimates with MACv2-SP (Table 2) is consistent with a too transparent atmosphere for shortwave radiation in climate models. Such a behaviour is typical for state-of-the-art radiation parameterisation schemes (e.g., Halthore et al., 2005; Randles et al., 2013) and has also been identified for the PSrad scheme (not shown) implemented in ECHAM and ECHAM-HAM. These and other reasons for model biases in anthropogenic aerosol forcing will be addressed in more detail in RFMIP, where models use the same MACv2-SP parameterisation as in the present study and accurate line-by-line radiation-transfer calculations will evaluate the necessary approximations for physical parameterisations in CMIP6 models.

In the following we evaluate the impact of observational uncertainty in τ_p and τ_a. For doing so, we assume the aerosol classification of MACv2 as an alternative representation (Offline-v2). 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 such that the MACv2 product primarily relies on global model results for separating anthropogenic from natural aerosols. Classifying additional fine-mode aerosol as anthropogenic as assumed by MACv2 increases the all-sky RF at TOA to -1.1 Wm^{-2}, which primarily arises due to stronger F_{aci} in MACv2. Ambiguous aerosol classifications, which occur especially for regions with low aerosol burden, and

a poor observational coverage are therefore reasons for uncertainty in present-day RF, i.e., the RF would be more negative if τ_a is assumed larger.

Choosing the larger anthropogenic fraction and lower background burden of MACv1 (Kinne et al., 2013), i.e., the year 1750 as a reference, when the background had less anthropogenic aerosol than 1850, yields a stronger RF in the offline model, namely an all-sky RF of -1.4 Wm^{-2} in SW at TOA (Offline-v1). Note that the clear-sky RF of the offline estimates and the complex models are in close 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 findings (Carslaw et al., 2013; Fiedler et al., 2017). Quantitative changes in natural aerosol burden between the pre-industrial and present-day remain unconstrained, e.g., model estimates of the anthropogenic fraction of desert dust are 10−60% associated with changes in land use and climate (Mahowald and Luo, 2003; Tegen et al., 2004; Stanelle et al., 2014). Since we cannot measure the aerosol of 1750 nor 1850, we propose using the present-day natural aerosol as background for a better comparability of observational and model estimates in future inter-comparison studies. Prescribing both the natural and anthropogenic aerosol across different models in future inter-comparison studies would allow to attribute remaining differences in the radiative effects to model errors in representing meteorological processes and radiative transfer.

3.4 Impact of spatial change of pollution

We assess the effect of a substantial spatial change of the τ_a maxima from Europe and the U.S. to East Asia between the mid-1970s and mid-2000s. One can additionally argue that the spatial differences in cloud regimes, insolation and surface albedo contribute to regionally different radiative effects resulting in a changing global ERF. For investigating this aspect we contrast the radiative forcing derived from the spatial distribution of the anthropogenic aerosol for the mid-1970s and mid-2000s (Figure 1). The different distribution of the anthropogenic aerosol clearly changes the pattern of their radiative effects (Figure 7). Namely, the maxima in regional contributions to RF and ERF occur over Europe and the U.S. in the mid-1970s, in contrast to the maximum over East Asia for the mid-2000s. The net contribution from adjustments is typically larger where the regional radiative forcings are largest in the mid-1970s.

Despite the regional differences in radiative effects and the inter-model spread in ensemble-averaged global mean RF and ERF, the spatial change of maximum aerosol pollution has little impact on the global mean RF and ERF of each model. The model ensemble mean changes from -0.59 Wm^{-2} for the mid-2000s to -0.54 Wm^{-2} for the mid-1970s. Likewise multi-year monthly means per model yield similar RFs for the two τ_a patterns (not shown). This implies that the seasonal contributions to RF are similar for both τ_a patterns, irrespectively which model we choose.

The ensemble-averaged change in ERF is small relative to natural year-to-year variability in modelled ERFs (Figure 8). Indeed, contrasting one-year estimates from the two aerosol patterns results in a large spread in ERF changes ranging from a decrease to an increase of ERF with the different τ_a patterns (Figure 8c−d). This is in line 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 any model. Moreover, it gives more evidence that the global mean ERF does not strongly depend on the northern hemispheric distribution of anthropogenic aerosol.

We better characterise the model behaviour for arriving at similar ERFs for the two τ_a patterns. For doing so, we calculate the regional forcing efficacies (E) for both RF and ERF in the shortwave at TOA, i.e., the ratio of the radiative effects and τ_a. We first average only over regions close to pollution sources ($\tau_a > 0.1$) and find that both E_{RF} and E_{ERF} are here stronger in the mid-2000s than for the mid-1970s for all models (Tables 3 and 4).

5 The behaviour of the models for E is, however, drastically different when we include areas further away from pollution sources. In this case, E_{RF} and E_{ERF} are typically stronger than close to pollution sources and have typically similar magnitudes for both aerosol patterns for each model, pointing to the importance of accurately knowing the spatial extent of aerosol pollution downwind. Of all models, NorESM and EC-Earth have the strongest E_{ERF} away from pollution sources indicating that the aerosol perturbation in those two models is more efficient in inducing radiative effects than in the other models. These two

10 models also show larger negative ERF than the other models. In EC-Earth this arises from the strongly negative net contribution from rapid adjustments, in NorESM from a strong negative RF combined with a small contribution from adjustments.

4 Conclusions

In the present work, we inter-compare ERFs from 180 years with annually repeating patterns of anthropogenic aerosol for each of the five state-of-the-art aerosol-climate models. The present-day all-sky ERF in the shortwave radiation at the top of

15 atmosphere is -0.59 Wm^{-2} using the multi-model, multi-member ensemble, where the anthropogenic aerosols are prescribed using the MACv2-SP parameterisation. The corresponding year-to-year standard deviation of 0.3 Wm^{-2} implies a typical year-to-year variability of 50%, reflecting both natural variability and model differences affecting ERF. We therefore propose a separation between long-term averaged ERF and estimates with super-imposed natural variability for studies on the ERF of anthropogenic aerosol. Based on the current work, we obtain a spread of -0.9 to -0.4 Wm^{-2} in the best model-mean estimates

20 of ERF$_{ari+aci}$, summarised in Fig. 9. In comparison to this model spread in ERF, the natural variability in ERF is large. For instance, the overall possible range in annual mean ERF$_{ari+aci}$ with superimposed natural variability is -1.5 to +0.5 Wm^{-2} in our multi-model ensemble. These differences in the spread underline the importance of using a sufficiently large number of years for quantifying ERF.

 Our results highlight that all models consistently show little change in the mean ERF of anthropogenic aerosol between

25 the mid-1970s and mid-2000s, despite the substantially different location of anthropogenic pollution maxima and the model diversity in their ERF magnitude. This is a remarkable result since the models run freely and differ in various model aspects including the representation of clouds and pre-industrial aerosol. Traditionally, such models have shown a substantial spread in ERF estimates (e.g., Shindell et al., 2013) comparable to the magnitude of the variability internal to the models in the present work. This behaviour suggests that diversity in anthropogenic aerosol optical properties, parameterising F_{aci}

30 in complex aerosol-climate models, and the large model-internal variability have a strong impact on ERF estimates. It gives further evidence that model-internal variability has not been sufficiently considered in past model inter-comparison studies tailored towards quantifying the model spread in ERF of anthropogenic aerosol, previously suggested based on ECHAM alone (Fiedler et al., 2017).

We recommend that studies on model differences in ERF consider the simulation length for evaluating whether model-internal variability has been sufficiently sampled, e.g., by using a confidence estimate (e.g., Fiedler et al., 2017). Note that natural variability is also an issue in constraining the magnitude of ERF from observations. Using the historical record of observations for constraining the ERF magnitude therefore should be done with ensembles of simulations or averaging over several decades. 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). Given our multi-model spread in absolute ERF magnitudes for the same τ_a, inter-comparing the relative changes in model-mean ERFs might herein give more stringent arguments for a model's value in representing the historical climate evolution. Our future work will focus on inter-comparing modelled ERF changes associated with other aerosol patterns for a better understanding of the historical evolution of ERF. One such endeavour is the usage of MACv2-SP in model simulations in the framework of CMIP6.

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 updated model version of the general circulation model that has been developed at MPI-M (Stevens et al., 2013). It is the atmospheric model of MPI-ESM1.2 participating in CMIP6 (Mauritsen et al., in review). ECHAM6.3 is a global hydrostatic model for the atmosphere with parameterisations of sub-grid scale physical processes. The atmospheric radiative transfer is parameterised with the PSrad scheme with the rapid radiative transfer model for general circulation models (RRTMG, Pincus and Stevens, 2013). External data sets define the boundary conditions of the model, including the climatology of surface properties, trace gas concentrations, and natural aerosol. A major change in MPI-ESM1.2 (Mauritsen et al., in review) compared to previous model versions is the implementation of MACv2-SP. The parameterisation prescribes anthropogenic aerosol optical properties and an associated Twomey effect (Fiedler et al., 2017; Stevens et al., 2017).

The global aerosol-climate model ECHAM6.3-HAM2.3 is an updated version of the one described by Tegen et al. (2018) and Neubauer et al. (submitted). Notable characteristics of this model version include updates on the atmospheric model and the sea-surface temperature dependent sea-salt emissions. 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 a two-moment stratiform cloud scheme (Lohmann and Hoose, 2009) for representing the activation of aerosol for forming cloud droplets and heterogeneously nucleating ice 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 dimethylsulphide (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). The prescribed background

aerosols are set to pre-industrial levels of HAM for all simulations. 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 and 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 Radiation Transfer Model (RRTM, 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 with aerosol emissions 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. Finally, new diagnostics have been added to IFS to allow calculation of instantaneous anthropogenic aerosol radiative effects using a double call of the radiation scheme (Section 2.2). 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 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 10 m 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 effects 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 aerosol 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 for ECHAM-HAM and ECHAM, NorESM sets all background aerosol emission to the values of 1850. These background conditions include sulphate from tropospheric volcanoes and from DMS, as well as organic matter from land 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). The pre-industrial burden in the aerosol-climate models contains some anthropogenic aerosol, but the majority of the pre-industrial aerosol optical depth (τ_p) is of natural origin. The 1850's global-mean τ_p in NorESM is namely 0.096, to which anthropogenic fossil-fuel aerosols contribute 0.002. For comparison, the year 2005 global-mean τ_a for MACv2-SP aerosols is 0.029.

Appendix B: Model diversity for clouds

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 models against observational climatologies from satellite products, listed in Table A1. The observational products herein provide an orientation for realistic values, although satellite retrievals also have caveats (e.g., Grosvenor et al., 2018).

We first assess the shortwave cloud radiative effect at the top of the atmosphere (F_{cld}), thus the cloud effect on the planetary albedo. The annual and global mean F_{cld} for $2001-2010$ from CERES Ed. 4 is -45.8 Wm^{-2}, i.e., less negative than that of 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} are generally speaking similar, but regional differences are distinct (Figure A1).

To better characterise the model diversity for clouds, we compare the global means in total cloud cover (f) and the ocean mean in vertically integrated liquid water content (l_{cld}) to satellite climatologies (Table A1). In the global mean, most models underestimate both f and l_{cld} compared to the satellite retrievals, but having too few clouds does not necessarily imply too little liquid or vice versa (Table A2). The spatial patterns (Figure A1) show a tendency of the models for underestimating f in the stratocumulus decks in the Southeastern regions of the Pacific and Atlantic Ocean where F_{aci} are thought to be important. The models, however, disagree on the regional values for f and l_{cld}. Moreover, the models show a large diversity in l_{cld} in the extra-tropical storm tracks. NorESM has herein the regionally largest l_{cld} exceeding 200 gm^{-2}. Our findings for l_{cld} are consistent with a similar regional evaluation of the relative differences of HadGEM and CAM (Malavelle et al., 2017), the latter of which has a similar atmospheric component to NorESM (see Appendix A).

The cloud differences raise the question how different the in-cloud droplet number concentration (N) for present day is in the models. We find that the prognostic schemes show large diversity in the pattern of N for present day, shown in Fig. A2. It is noteworthy that N from the prognostic schemes is for stratiform cloud types, but can additionally include detrained N from anvils of deep convection. The spatial pattern of N in ECHAM is not shown due to the simplistic treatment in the model.

5   ECHAM has statically prescribed values for N that are constant with height below 800 hPa and exponentially decrease aloft. The near-surface values in ECHAM are $N = 80\,\mathrm{cm}^{-3}$ over ocean and $N = 180\,\mathrm{cm}^{-3}$ elsewhere (not shown), and are multiplied with η_N from MACv2-SP like in the other models.

Compared to the satellite product, models with prognostic schemes 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

10   models and the satellite product is due to differences in the methods for diagnosing N in the satellite and model approaches, 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 an artefact of compensating differences from tuning the radiation balance at the top of the atmosphere in the models. For instance, the behaviour of NorESM points to too much shortwave reflectivity by too thick clouds that overcompensate the missing reflection

15   due to underestimated cloud cover.

Author contributions. SF designed the study, performed the experiments with ECHAM, analysed the data of all models, and lead the writing of the manuscript. SK performed the offline radiative transfer calculation. PR performed the experiments with NorESM, KH for ECHAM-HAM, NB for HadGEM, and TvN and DOD for EC-Earth. All authors contributed to the discussion of the results and the writing of the manuscript.

20   *Competing interests.* The authors confirm that they have no competing interests.

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.

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. and 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., 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.

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, submitted.

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. and et al.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and its response to increasing CO2, in review.

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,

17

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.

5 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.

10 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.,
15 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,
20 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
25 in the atmosphere, Journal of Geophysical Research: Atmospheres, 116, doi:10.1029/2010JD014713, 2011.

Stanelle, T., Bey, I., Raddatz, T., Reick, C., and Tegen, I.: Anthropogenically induced changes in twentieth century mineral dust burden and the associated impact on radiative forcing, Journal of Geophysical Research: Atmospheres, 119, 13,526–13,546, doi:10.1002/2014JD022062, 2014.

Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., Salzmann, M., Schmidt, H., Bader, J., Block, K., Brokopf, R.,
30 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.

35 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., 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., Kirkevåg, 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., Werner, M., Harrison, S. P., and Kohfeld, K. E.: Relative importance of climate and land use in determining present and future global soil dust emission, Geophysical Research Letters, 31, doi:10.1029/2003GL019216, 2004.

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, 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.

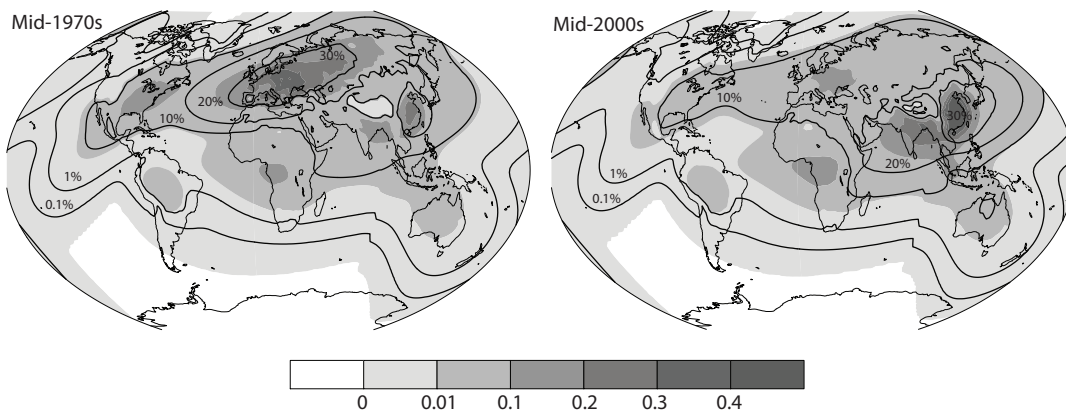**Figure 1.** Mean anthropogenic aerosol optical depth (τ_a, shaded) and fractional increase in cloud droplet number (η_N, contours) associated with anthropogenic aerosol. Shown are annual means of τ_a at 550nm and η_N for the (left) mid-1970s and (right) mid-2000s from MACv2-SP that prescribes annually repeating monthly maps of τ_a in the participating models. Note the non-linear scale for also displaying small values.
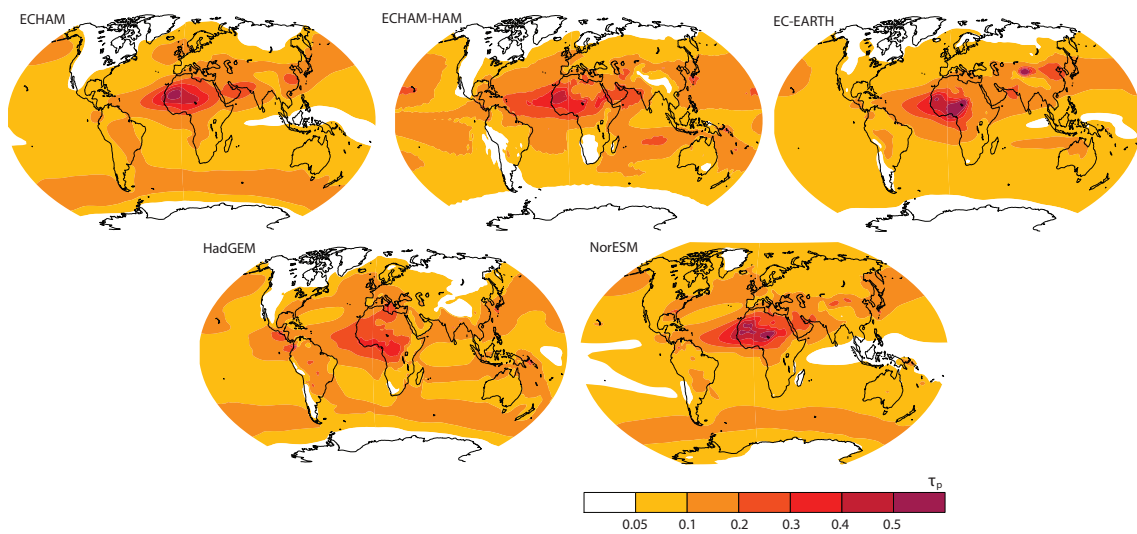
Figure 2. Mean pre-industrial AOD. Shown are annual means of τ_p of the radiation band around 550 nm for each model.
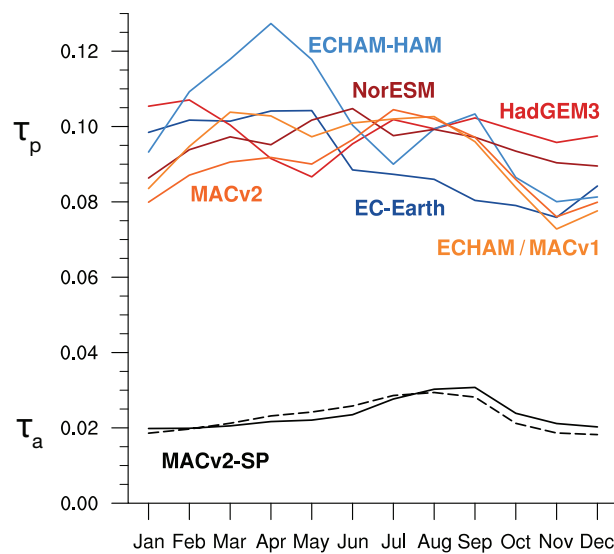
Figure 3. Annual cycle of the global mean AOD at 550nm. Shown are monthly means of (colors) τ_p from the models and (black) τ_a for the (dashed) mid-1970s and (solid) mid-2000s from MACv2-SP.
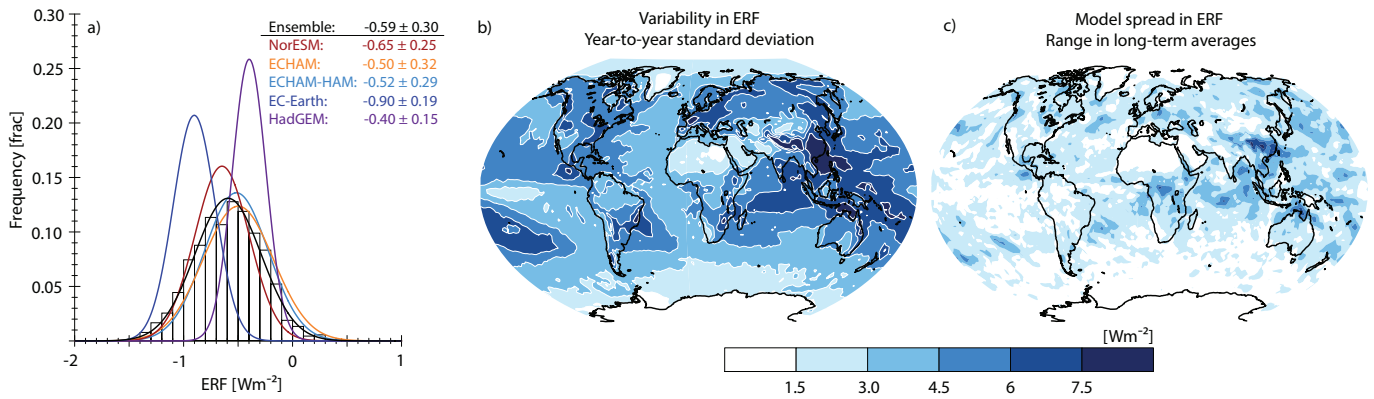
Figure 4. Variability in annual ERF estimates for the mid-2000s. Shown are (a) Gaussian distributions of annual ERF estimates for present-day from (colors) individual model ensembles and (black) the entire multi-model, multi-member ensemble, (b) the regional standard deviation of annual contributions to ERF from the entire multi-model, multi-member ensemble as measure for the variability internal to the model ensemble, and (c) the range in the long-term averaged ERFs of the models as measure for the spread in ERF associated with model differences. In (a), 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. ERF values are for SW at the TOA for all-sky conditions.
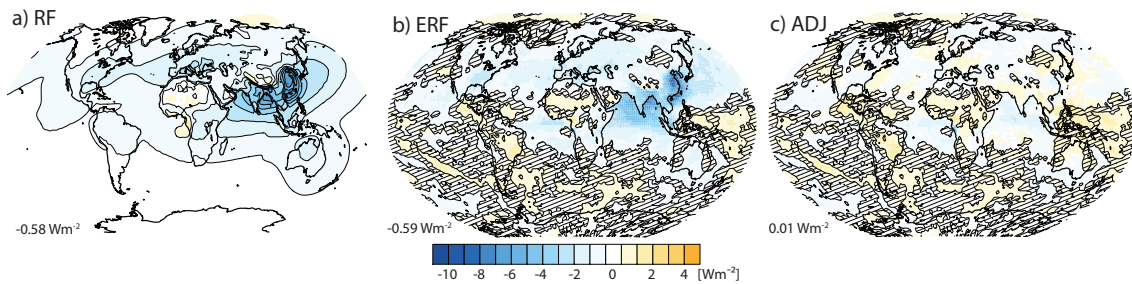
a) RF

-0.58 Wm⁻²

b) ERF

-0.59 Wm⁻²

c) ADJ

0.01 Wm⁻²

-10 -8 -6 -4 -2 0 2 4 [Wm⁻²]

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 ERF at a 10% significance level. The numbers in the lower left corner are the spatial averages.
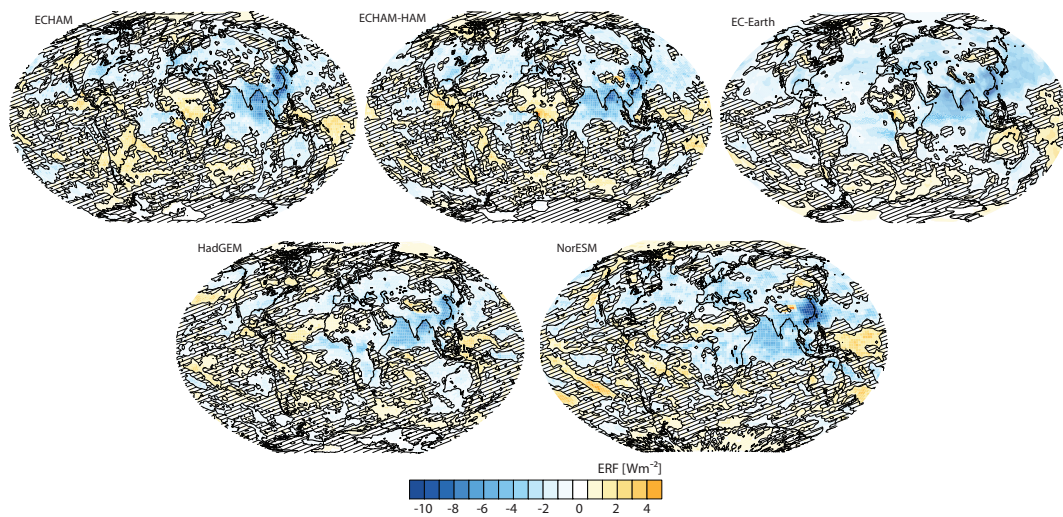
Figure 6. Multi-member ensemble mean of effective radiative effects of anthropogenic aerosol for the mid-2000s. Shown are the effective radiative forcing for SW at the TOA in all-sky conditions for each model. Hatching indicates non-significant ERF at a 10% significance level.
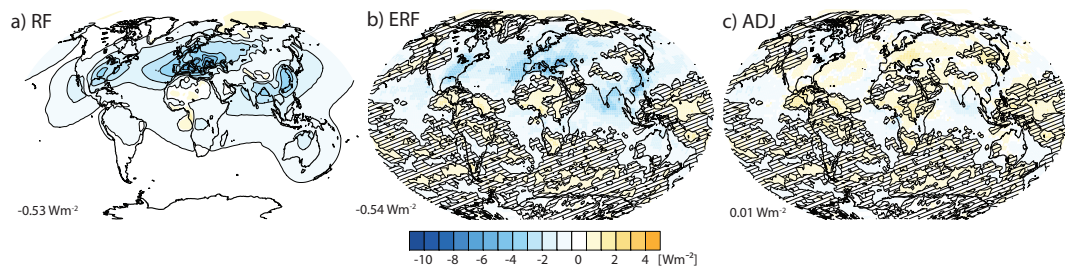
Figure 7. Multi-model, multi-member ensemble mean of the anthropogenic aerosol radiative effects for the mid-1970s. As Figure 5, but with the anthropogenic aerosol pattern of the mid-1970s.
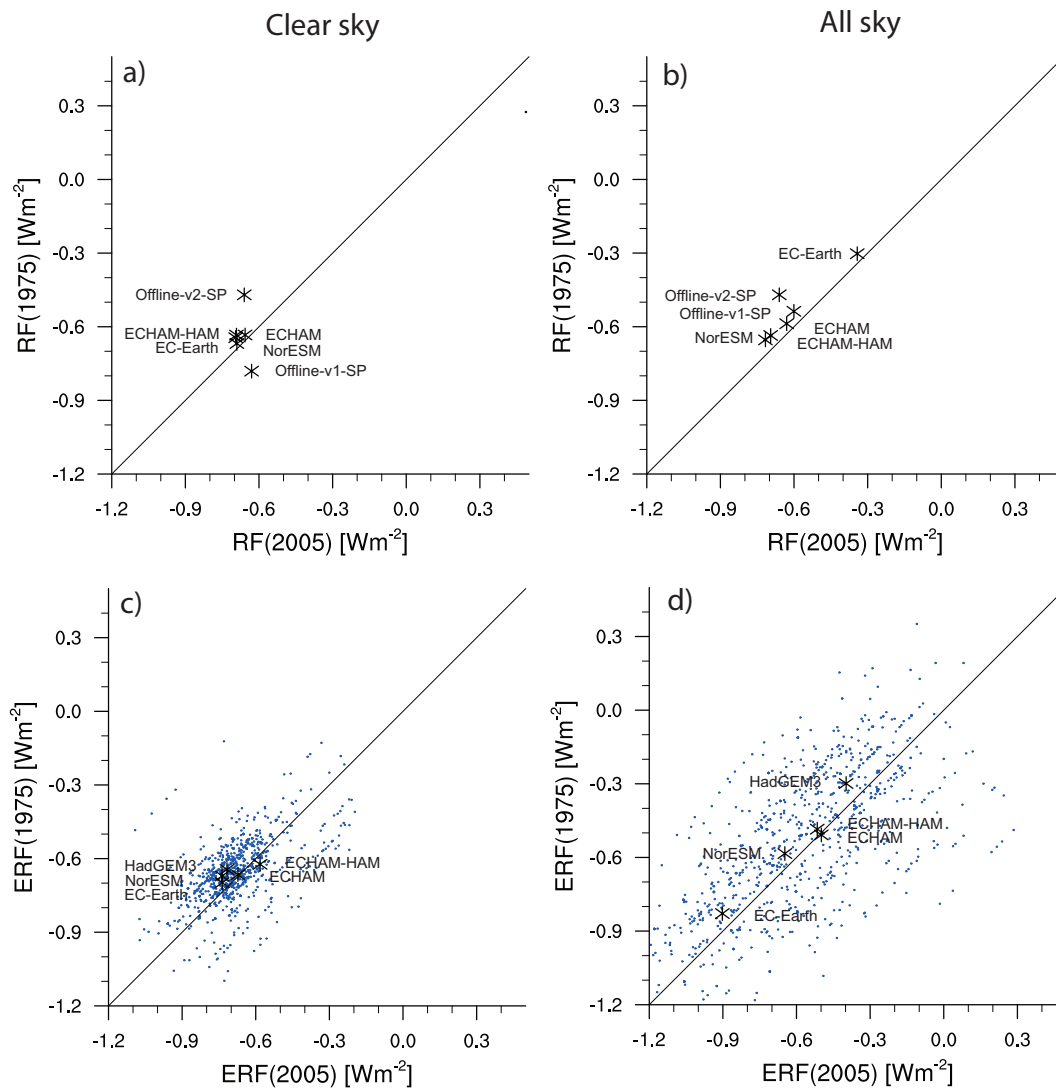
Figure 8. Anthropogenic aerosol forcing of the mid-1970s against the mid-2000s. Shown are the (top) instantaneous and (bottom) effective radiative forcing for SW at the TOA from the pollution of the mid-1970s against the mid-2000s for (left) clear and (right) all sky. Thick crosses are the ensemble means. Blue dots in (c, d) are the model averages of individual years representing the year-to-year variability internal to the model ensemble.
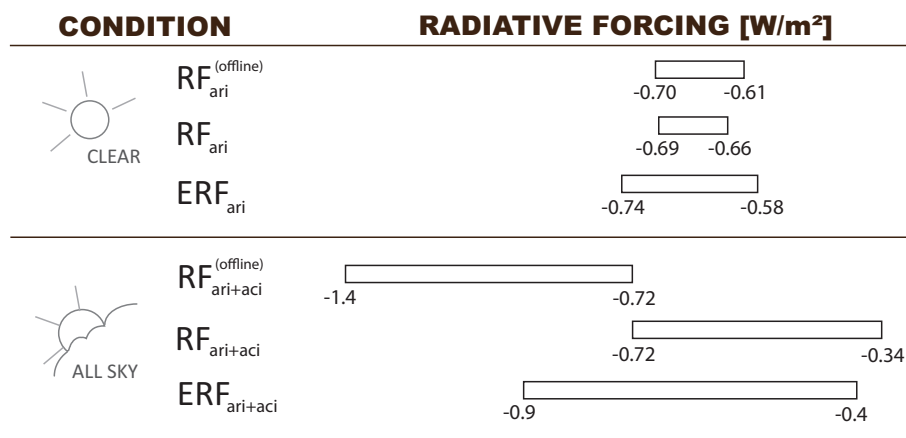
CONDITION	RADIATIVE FORCING [W/m²]

CLEAR

$RF_{ari}^{(offline)}$ — -0.70 -0.61

RF_{ari} — -0.69 -0.66

ERF_{ari} — -0.74 -0.58

ALL SKY

$RF_{ari+aci}^{(offline)}$ — -1.4 -0.72

$RF_{ari+aci}$ — -0.72 -0.34

$ERF_{ari+aci}$ — -0.9 -0.4

Figure 9. Overview on model spread in anthropogenic aerosol forcing for the mid-2000s. Shown are the instantaneous (RF) and effective radiative forcing (ERF) associate with aerosol-radiation and aerosol-cloud interaction for SW at the TOA for clear and all sky from Tab. 2. RF from the offline calculations consider additional uncertainty sources and are shown as separate bars. Refer to Section 2.1 for details.

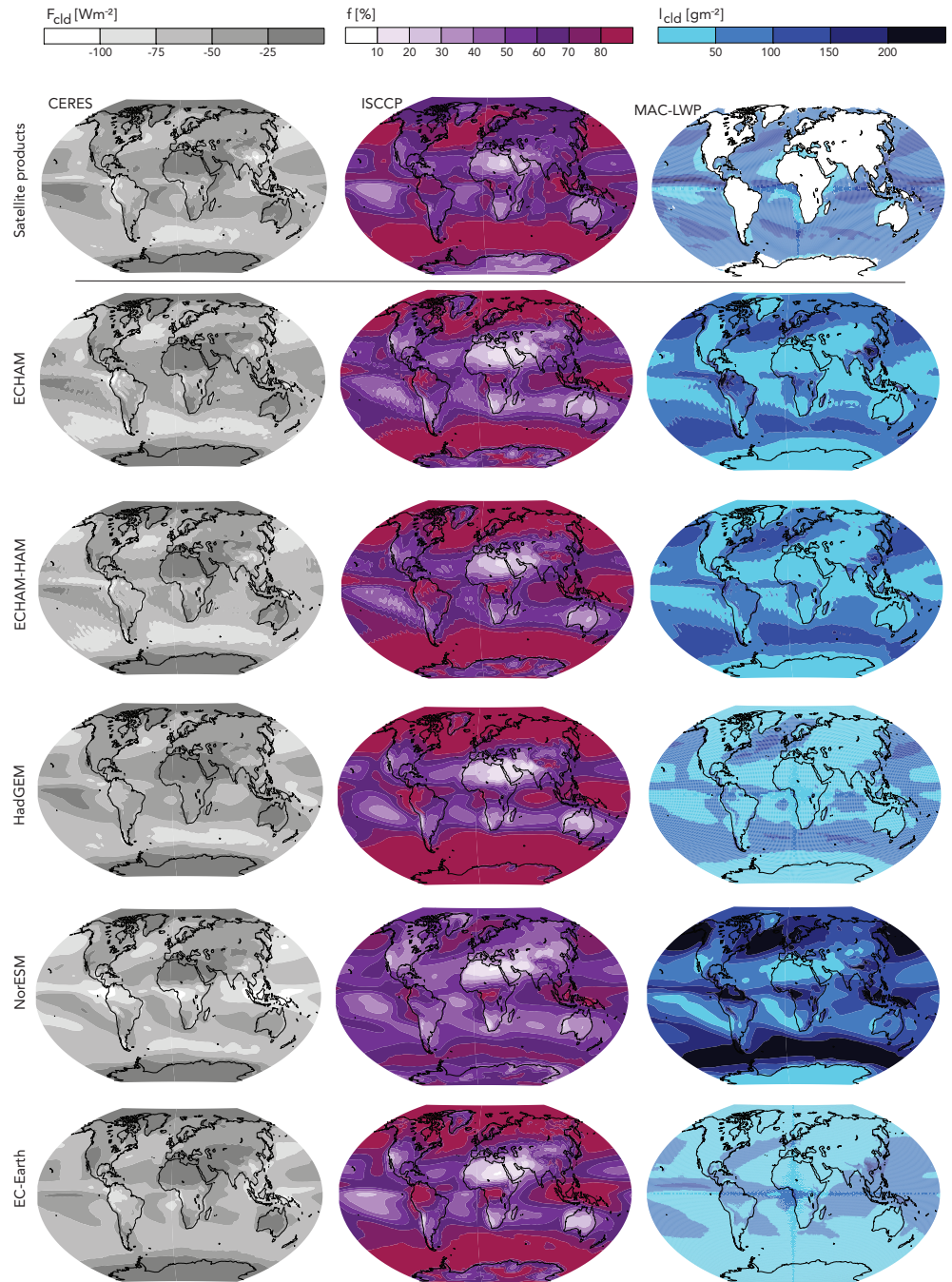**Figure A1.** Multi-member ensemble means of cloud characteristics for the mid-2000s. Shown are the mean (left column) SW cloud radiative effect at the TOA, F_{cld}, (middle column) total cloud cover, f, and (right column) vertically integrated liquid water content, l_{cld} from (top row) satellite products and (rows beneath) the models (Table A1). Areas without reliable satellite retrieval are shaded white.
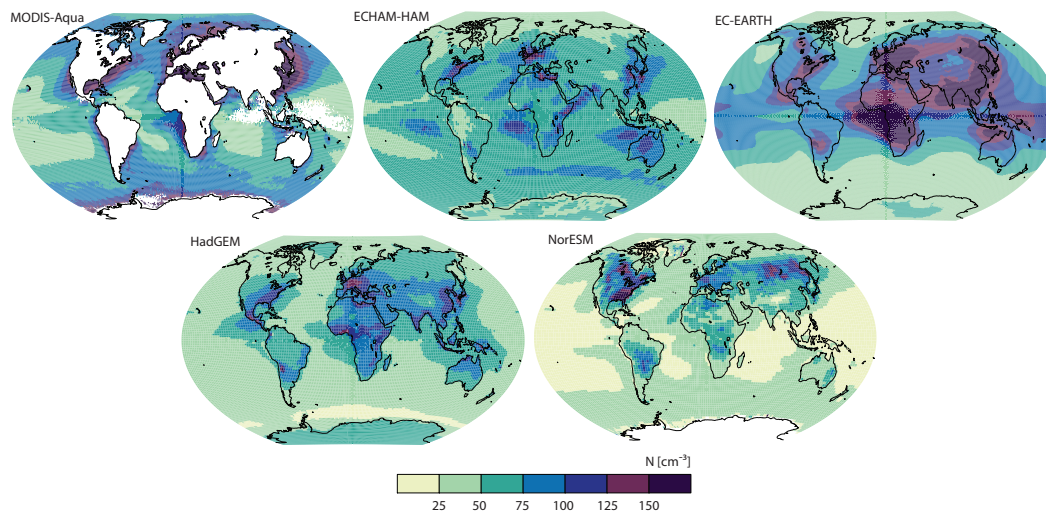
Figure A2. In-cloud droplet number concentration for mid-2000s. Shown are the annually and vertically averaged in-cloud droplet number concentration (N) from the aerosol-climate models and from the MODIS-Aqua satellite product by Bennartz and Rausch (2017). Areas without reliable satellite retrieval are shaded white.

Table 1. Model experiment setup

Model	Δx x Δy	Levels	τ_p	τ_a
ECHAM	1.875°E x 1.875°N	47	MACv1 clim	MACv2-SP
ECHAM-HAM	1.875°E x 1.875°N	47	Online	MACv2-SP
EC-Earth	1.875°E x 1.875°N	91	TM5 clim	MACv2-SP
HadGEM3	1.875°E x 1.25°N	85	HadGEM3 clim	MACv2-SP
NorESM	2.5°E x 1.894°N	26	Online	MACv2-SP
Offline-v2-SP	1°E x 1°N	20	MACv2	MACv2-SP
Offline-v1-SP	1°E x 1°N	20	MACv1	MACv2-SP
Offline-v2	1°E x 1°N	20	MACv2	MACv2
Offline-v1	1°E x 1°N	20	MACv1	MACv1

Table 2. Ensemble averages of instantaneous (RF) and effective (ERF) radiative forcing, and net contribution from rapid adjustments (ADJ) at the surface and the TOA in SW for all sky (clear sky) in Wm^{-2} for τ_a of 2005. The first block shows aerosol-climate models with MACv2-SP, and the second block shows offline benchmarks.

	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	-1.34 (-1.81)	-0.34 (-0.69)	-0.90 (-0.74)	-0.6 (0.05)
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 3. Ensemble averages of regional forcing efficacies (E) for the mid-2000s at the TOA in SW for all sky (clear sky) in Wm^{-2}. E is calculated as RF or ERF divided by τ_a and spatially averaged over regions either near pollution sources or additionally areas further away, i.e., $\tau_a > 0.1$ and $\tau_a > 0.01$ (Figure 1).

	$E_{RF}(\tau_a > 0.01)$	$E_{RF}(\tau_a > 0.1)$	$E_{ERF}(\tau_a > 0.01)$	$E_{ERF}(\tau_a > 0.1)$
ECHAM	-26 (-24)	-17 (-25)	-22 (-26)	-18 (-25)
ECHAM-HAM	-32 (-26)	-20 (-26)	-22 (-21)	-15 (-25)
EC-Earth	-12 (-25)	-13 (-27)	-41 (-29)	-21 (-26)
HadGEM3	/	/	-11 (-27)	-16 (-26)
NorESM	-30 (-24)	-19 (-27)	-31 (-27)	-21 (-28)

Table 4. As Table 3, but for the mid-1970s.

	$E_{RF}(\tau_a > 0.01)$	$E_{RF}(\tau_a > 0.1)$	$E_{ERF}(\tau_a > 0.01)$	$E_{ERF}(\tau_a > 0.1)$
ECHAM	-26 (-27)	-11 (-18)	-23 (-29)	-13 (-18)
ECHAM-HAM	-32 (-29)	-13 (-19)	-20 (-27)	-9 (-18)
EC-Earth	-13 (-27)	-8 (-18)	-44 (-29)	-16 (-19)
HadGEM3	/	/	-8 (-28)	-10 (-18)
NorESM	-31 (-27)	-14 (-19)	-30 (-29)	-13 (-20)

Table A1. Gridded satellite climatologies as reference.

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)	F_{cld} [Wm^{-2}] Cloud shortwave radiative effects at top of the atmosphere	2001−2014
ISCCP	International Satellite Cloud Climatology Project (Rossow and Schiffer, 1999)	f [%] Total cloud cover	1983−2009
MAC-LWP	Multi-sensor Advanced Climatology (Elsaesser et al., 2017)	l_{cld} [gm^{-2}] Liquid water path	2000−2016
MODIS	Climatology based on Moderate Resolution Imaging Spectroradiometer aboard Aqua (Bennartz and Rausch, 2017)	N [cm^{-3}] Warm cloud droplet number concentration	2003−2015

Table A2. Global mean statistics for clouds and aerosol. l_{cld} and N are herein averages over ocean regions, consistent with the satellite data availability (Figures A1 and A2). The details on the satellite products are listed in Tab. A1.

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