Rethinking the Lower Bound on Aerosol Radiative Forcing

BJORN STEVENS

Max Planck Institute for Meteorology, Hamburg, Germany

(Manuscript received 25 September 2014, in final form 27 February 2015)

ABSTRACT

Based on research showing that in the case of a strong aerosol forcing, this forcing establishes itself early in the historical record, a simple model is constructed to explore the implications of a strongly negative aerosol forcing on the early (pre-1950) part of the instrumental record. This model, which contains terms representing both aerosol–radiation and aerosol–cloud interactions, well represents the known time history of aerosol radiative forcing as well as the effect of the natural state on the strength of aerosol forcing. Model parameters, randomly drawn to represent uncertainty in understanding, demonstrate that a forcing more negative than $-1.0\,\mathrm{W\,m^{-2}}$ is implausible, as it implies that none of the approximately 0.3-K temperature rise between 1850 and 1950 can be attributed to Northern Hemisphere forcing. The individual terms of the model are interpreted in light of comprehensive modeling, constraints from observations, and physical understanding to provide further support for the less negative ($-1.0\,\mathrm{W\,m^{-2}}$) lower bound. These findings suggest that aerosol radiative forcing is less negative and more certain than is commonly believed.

1. Introduction

A perturbation to the composition of Earth's atmosphere can be quantified through the degree to which it disturbs the radiative balance at the top of the atmosphere, its radiative forcing. This radiative forcing is a motive force for climate change as (at least for small perturbations) Earth's globally averaged surface temperature is expected to change proportionally with the forcing (e.g., Myhre et al. 2013a; Sherwood et al. 2015). More than 20 years ago Charlson et al. (1992) used simple physical arguments to raise the specter of a relatively large but negative ($-2.3\,\mathrm{W\,m^{-2}}$) radiative forcing by tropospheric aerosols resulting from human activities. Although subsequent assessments (e.g., Boucher et al. 2013) have suggested that the present-day radiative forcing by the tropospheric aerosol ($\overline{F}_{\mathrm{aer}}$) is somewhat smaller ($-0.9\,\mathrm{W\,m^{-2}}$), the specter of a large forcing lingers as uncertainty (ranging from -0.1 to $-1.9\,\mathrm{W\,m^{-2}}$ for a 90% confidence interval) arises from a poor understanding of how clouds respond to aerosol perturbations.

One important implication of a strongly negative aerosol forcing is that Earth's globally averaged surface temperature must be very sensitive to greenhouse gas forcing to have risen at all over the instrumental record. Another implication is that if $\overline{F}_{\mathrm{aer}}$ fails to intensify apace with the positive greenhouse gas forcing, for instance because of efforts to reduce pollution, Earth's surface temperatures will rise more rapidly (Charlson et al. 1991; Brasseur and Roeckner 2005).

The complexity of the processes leading to an aerosol forcing is daunting, and understanding remains rudimentary. The scale of the processes controlling the lifetime and composition of the aerosols and their interaction with clouds is far below what can be resolved by a large-scale model. So even if these processes were well understood, it would be far from trivial to represent with any quantitative fidelity their collective effects on the scales of motion representable by a global model. So it is not surprising that Earth system models constructed to estimate $\overline{F}_{\mathrm{aer}}$ through an incorporation of aerosol processes remain sensitive to a large number of poorly constrained assumptions (Boucher et al. 2013; Hoose

Denotes Open Access content.

Corresponding author address: Bjorn Stevens, Max Planck Institute for Meteorology, Bundesstrasse 53, 20146 Hamburg, Germany.
E-mail: bjorn.stevens@mpimet.mpg.de

DOI: 10.1175/JCLI-D-14-00656.1

et al. 2009; Golaz et al. 2011). Even in the modern period, during which aerosols have been relatively well observed by networks of ground stations, advanced surface remote sensing, numerous airborne measurements, and a constellation of satellite sensors, Earth system models do not agree as to whether radiative forcing[1] has been increasing, decreasing, or not changing at all (Shindell et al. 2013; Kühn et al. 2014; Carslaw et al. 2013).

For these reasons, I believe there is little foundation for the expectation that comprehensive modeling alone can provide a basis for reducing uncertainty in estimates of $\overline{F}_{\text{aer}}$ or that somehow these models encapsulate uncertainty in understanding. Fortunately, even for very complex problems, simple-minded approaches, when targeted to a particular aspect of the problem, can sometimes provide surprising insights. The present lower bound for $\overline{F}_{\text{aer}}$ is a case in point, as it stems from the realization that a more negative forcing would be incompatible with the observational record since 1950 (Murphy et al. 2009). In this paper I argue similarly, but instead combine an understanding of the temperature record before 1950 with physical reasoning and insights arising from the robust response of comprehensive models to make the case for a substantially less negative lower bound for $\overline{F}_{\text{aer}}$ ($-1.0\,\text{W}\,\text{m}^{-2}$). When combined with an upper bound derived from the more recent observational record (Murphy et al. 2009) this implies an uncertainty range in aerosol forcing between -0.3 and $-1.0\,\text{W}\,\text{m}^{-2}$, reducing by nearly a factor of 3 the uncertainty in this important quantity.

My main arguments are developed in three parts. First, I develop and motivate a simple model designed to represent the time history of $\overline{F}_{\text{aer}}$, and through which a revised lower bound of its present-day value is derived. Second, I interpret this model in light of present understanding of aerosol processes, and show that this understanding is consistent with a less negative lower bound on $\overline{F}_{\text{aer}}$. Third, through an analysis of simulations conducted as part of phase 5 of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012), I argue that a smaller forcing is implied when one compares the response of the models to available observations. At the end of the manuscript the implications of my findings are discussed. In a series of appendixes further theoretical justification is given for the simple model used to interpret the historical aerosol forcing. In addition, the methods, models, and a more complete description of primary data sources are presented.

2. A simple model for the time history of aerosol forcing

The central idea developed in this paper is that if aerosol forcing arising from the interactions between aerosols and clouds increases sublinearly with emissions, for instance logarithmically as is suggested both by physical understanding and comprehensive modeling (Charlson et al. 1992; Carslaw et al. 2013), then a disproportionate amount of the forcing would be expected to arise early in the instrumental record. Put another way, one unit of emissions in a pristine atmosphere can be expected to introduce a larger radiative forcing than one unit of emissions in an atmosphere already burdened by substantial anthropogenic emissions. The implication is that during the early part of the industrial period aerosol forcing will have increased disproportionately compared to greenhouse gas forcing, and hence it might be informative to look to this period to help disentangle the effect of the radiative forcing of aerosols from other anthropogenic forcings.

To take advantage of this line of thought one requires a model capable of resolving the temporal evolution of aerosol forcing. In principle a climate–chemistry model could be used for this purpose. In practice the most comprehensive models are usually run for short time periods with preindustrial climate forcings, and again for the present day, with relatively little regard to what happens in between. Simulating the entire history of the industrial period with a comprehensive model is computationally expensive, but not prohibitively so; but assessing and sampling the uncertainty space of such a model is another story (cf. Carslaw et al. 2013). To circumvent this difficulty, I posit a functional form for the aerosol forcing whose time dependence is carried solely by the global emission history of sulfur dioxide (SO_2), which I denote by $\overline{Q}_a$.

Parameterizing $\overline{F}_{\text{aer}}$ as a function of $\overline{Q}_a$ has a long history, most notably dating back to the seminal work of Charlson et al. (1992). For readers unfamiliar with such an approach it may be helpful to review the material in appendixes A and B where the physical justification for relating $\overline{F}_{\text{aer}}$ to $\overline{Q}_a$ is developed more systematically. Parameterizing $\overline{F}_{\text{aer}}$ as a function of $\overline{Q}_a$ is attractive because emissions of SO_2 from the combustion of fossil fuels, biomass, and metal smelting are reasonably well known as summarized by Smith et al. (2011) (see Fig. 1). Because emissions are bounded by the available sulfur in the fuel

[1] Here and throughout radiative forcing defined as a global quantity. Although there is clear evidence of regional changes in emissions of aerosols and aerosol precursors during the modern period, the available evidences suggests that these have at most a regional imprint (e.g., Murphy et al. 2009; Stevens and Schwartz 2012; Murphy 2013; Bengtsson and Schwartz 2013).

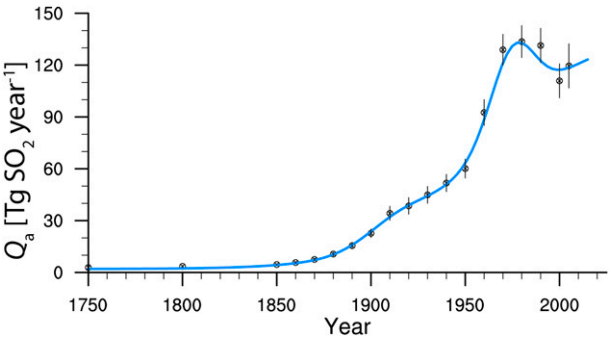

FIG. 1. Anthropogenic SO_2 emissions $\overline{Q}_a$ from Smith et al. (2011). By 1950 SO_2 emissions had reached half their present-day value. The blue line is a fit to the data (see appendix C).

(or ore) they are relatively well constrained[2] as compared to some other by-products of combustion (Smith et al. 2011).

In the remainder of this section the two main premises of the simple model for $\overline{F}_{aer}$ are outlined in more detail, after which the implications of aerosol forcing varying as predicted by the simple model are explored and from which a new lower bound on aerosol forcing is derived.

a. *Scaling aerosol forcing by SO_2 emissions*

As a first approximation I simply assume that $\overline{F}_{aer}$ can be expressed as a function of $\overline{Q}_a$ As an empirical statement my assumption is well supported by available data and modeling as illustrated by Fig. 2.

To demonstrate the form of the relationship between $\overline{F}_{aer}$ and $\overline{Q}_a$ in Fig. 2, three sources of information have been compiled and normalized: calculations from the model intercomparison study of Shindell et al. (2013) are shown as the black-filled circles, calculations from the modeling study of Carslaw et al. (2013) are shown by the blue-filled circles, and values tabulated in Annex II of the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5) (Prather et al. 2013) are shown as smaller gray-filled circles. Because Carslaw et al. only estimate the forcing from aerosol–cloud interactions ($\overline{F}_{aci}$), a contribution from aerosol–radiation interactions ($\overline{F}_{ari}$) is added to those authors forcing estimates to derive an estimate of the total aerosol forcing. The contribution by $\overline{F}_{ari}$, which is added to the Carslaw et al. estimate of $\overline{F}_{aci}$, is calculated by assuming that $\overline{F}_{ari}$ scales linearly with $\overline{Q}_a$ (for reasons justified later) and by setting the present-day value of $\overline{F}_{ari}$ to $-0.45\,W\,m^{-2}$ following IPCC AR5 (Boucher et al. 2013). For the data points taken from the Shindell et al. (2013) and Carslaw et al. (2013) studies, values of $\overline{F}_{aer}$ are linearly scaled to yield $\overline{F}_{aer} = -0.9\,W\,m^{-2}$ for the present day, in agreement with data from Annex II of AR5. To the extent that the data fall on one curve, it suggests that although different models might disagree on the sensitivity of $\overline{F}_{aer}$ to $\overline{Q}_a$, there is some robustness in the *form* of the relationship. The data collapse reasonably well, especially over the period between 1900 and 1980 when anthropogenic burdens increase the most.

A critical eye might complain that after 1980 and prior to 1900 the form of the relationship between $\overline{F}_{aer}$ and $\overline{Q}_a$ shows more dependence on the source of the estimate. For instance, after 1980 some estimates show $\overline{F}_{aer}$ changing more than would be expected given the relatively small change in $\overline{Q}_a$. And before 1850, the estimate for $\overline{F}_{aer}$ from the AR5 is much more sensitive to changes in $\overline{Q}_a$ than is the estimate by Carslaw et al. (2013). But the figure also shows that the post-1980 departures from the midcentury form of the relationship between $\overline{F}_{aer}$ and $\overline{Q}_a$ are not robust, and that for the early period the AR5 estimates are not believable. The AR5 estimate of $\overline{F}_{aer}$ implies a tremendous sensitivity to $\overline{Q}_a$ prior to 1850, relatively small sensitivity through the first part of the twentieth century, and again a larger sensitivity between 1950 and 1980. There is no real physical basis for such strong and discrete changes with time, particularly during that period. Although changes in fuels or methods of combustion could affect aerosol optical properties, for instance through the coemission of other aerosol precursors and/or black carbon, and the ice-core record provides some evidence of this (McConnell et al. 2007; Fischer et al. 1998), the observed changes cannot explain shifts in the AR5 record. More plausibly they result from the effects of using different modeling studies for different time periods when constructing the AR5 estimate (Shindell et al. 2013).

Nonetheless, because $\overline{F}_{aer}(\overline{Q}_a)$ relates the globally averaged forcing to globally averaged sources, it stands to reason that even if aerosol radiative effects scale with the local burden, changes in the distribution and nature of sources will change the relationship between $\overline{F}_{aer}$ and $\overline{Q}_a$ over time. This type of effect is expected to be most pronounced for $\overline{F}_{aci}$, which saturates as burdens increase, thereby increasing the sensitivity of the forcing to the spatial and temporal distribution of the burden.

To the extent that changing patterns of emissions are important for the global forcing, it would be more appropriate to express $\overline{F}_{aer}$ as a function of the source strength of the different patterns of emissions, something that comprehensive models are designed to do.

[2] Carslaw et al. (2013) estimate that uncertainty in $\overline{Q}_a$ makes a nonnegligible contribution (38%) to the forcing uncertainty; however, their quantification was based on an informal estimate (expert survey) of uncertainty that was a factor of 4 larger than that estimated by Smith et al. (2011).

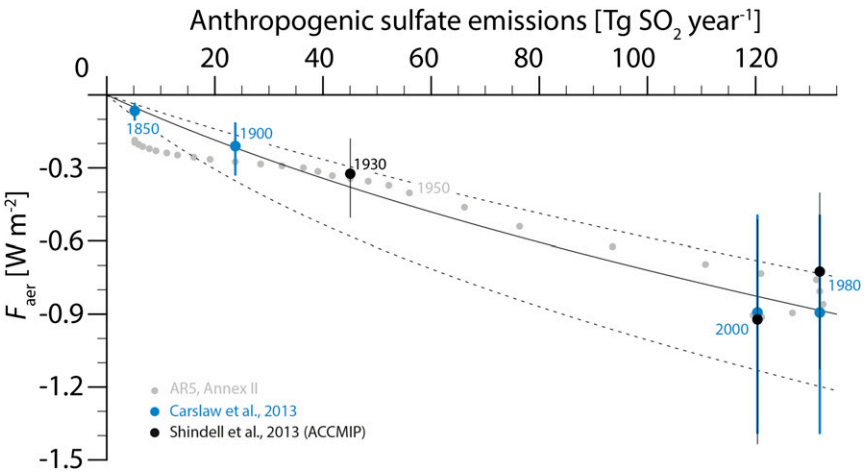

FIG. 2. Different estimates of the aerosol forcing, normalized to the best estimate of the IPCC AR5, vs anthropogenic sulfate emissions taken from Smith et al. (2011). Estimates from comprehensive modeling are provided for recently published studies (Carslaw et al. 2013; Shindell et al. 2013) by large filled circles with 1σ uncertainty indicated by the vertical lines. For comparison to the AR5, all estimates are normalized to correspond to the AR5 best estimate, as in the AR5 Annex II (see text for further details). AR5 time series of aerosol forcing are shown by the closed gray circles. The solid line through the data is based on Eq. (1) with $\alpha = 0.001\,875$, $\beta = 0.634$, and $\overline{Q}_n = 76$ Tg SO_2 yr^{-1}. Following work that shows that a stronger aerosol forcing is associated with a weaker natural sulfate source (Carslaw et al. 2013). The dotted lines show how $\overline{F}_{aer}$ changes if $\overline{Q}_n$ is increased or decreased by 50%.

Two of the three models [Geophysical Fluid Dynamics Laboratory Atmospheric Model, version 3 (GFDL AM3), and GISS-E2-R; expansions of model names and other acronyms and abbreviations are available online at http://www.ametsoc.org/PubsAcronymList] analyzed by Shindell et al. (2013) for the period between 1980 and the present day indeed show that, starting in the 1990s, a multiple (rather than single) pattern-based approach might be necessary to encapsulate the global forcing, as the rise of SO_2 emissions in South and East Asia give rise to a forcing from aerosol–cloud interactions that more than offsets the reduction in forcing caused by declining North American and European emissions. The response of these two models explains the scatter in the comprehensive modeling estimates at high sulfate burdens in Fig. 2 and is the basis of the claim by Shindell et al. (2013) that, despite a reduction in $\overline{Q}_a$, $\overline{F}_{aer}$ has become more negative over the past 30 years. However, the signal underlying this claim is very small compared to the uncertainties in the modeling, and is not robust; an equal number of studies show no change in forcing between 1980 and 2000 [e.g., the blue points in Fig. 2, which are taken from Carslaw et al. (2103), as well as results from the CSIRO model, which was the third one analyzed by Shindell et al.]. A more recent study even shows that there is a strong decrease in the magnitude of $\overline{F}_{aer}$ over the same period (Kühn et al. 2014).

The statement that $\overline{F}_{aer}$ can be expressed as a function of $\overline{Q}_a$ has more than just empirical support. As mentioned at the top of this section, there are good physical justifications for expressing $\overline{F}_{aer}$ as a function of $\overline{Q}_a$. These are discussed in more detail in section 4 and appendixes A and B and motivate the development of the physical model introduced in the next section. But in simpler terms, the strength of the relationship in Fig. 2 can be interpreted as a manifestation of the idea that the net forcing is proportional to the globally averaged sulfate burden, and that sulfate burdens are proportional to $\overline{Q}_a$. The first point follows either because the radiative forcing from the sulfate aerosol still dominates the total forcing or because the anthropogenic burdens of important nonsulfate aerosols are reasonably well correlated with sulfate burdens. A correlation between anthropogenic sulfate burdens and burdens of nonsulfate anthropogenic aerosols could arise because they are coemitted or simply because sulfate burdens are a good indicator of human activity. In either case the strong relationship between $\overline{F}_{aer}$ and $\overline{Q}_a$ is indicative of the fact that differences associated with changing patterns of emissions and changes in the mix of nonsulfate aerosols do not project strongly on to the time history of $\overline{F}_{aer}$. The second point follows from the idea that changes in the oxidation rate of SO_2 and sulfate lifetimes are small compared to changes in $\overline{Q}_a$.

In summary, the idea that $\overline{F}_{\rm aer}$ scales with $\overline{Q}_a$, long a linchpin of arguments for a strong aerosol forcing (e.g., Charlson et al. 1992), remains a reasonable assumption for exploring the variation of global forcing over the historical period as a whole. To put it another way, although one can certainly imagine why $\overline{F}_{\rm aer}$ need not remain a simple function of $\overline{Q}_a$, present understanding does not warrant abandoning the considerable simplification that this assumption entails.

b. A simple model

Having established that expressing $\overline{F}_{\rm aer}$ a function of $\overline{Q}_a$ is a reasonable assumption, the form of this relationship remains to be determined. I propose that

$$\overline{F}_{\rm aer} = -\alpha\overline{Q}_a - \beta \ln\left(\frac{\overline{Q}_a}{\overline{Q}_n} + 1\right). \qquad (1)$$

The time dependence of this expression is carried by the time dependence of $\overline{Q}_a$. A natural source $\overline{Q}_n$ is included in the model, which can be interpreted as the equivalent magnitude of the SO_2 source required to produce the observed natural distribution of cloud droplets. That along with α and β constitute the sole model parameters. The simplicity of Eq. (1) belies the amount of research that underpins it, a justification that I outline briefly below and elaborate upon in the appendixes. The arguments build very much on those introduced by Charlson et al. (1992) but the treatment of individual terms builds on what we have learned since that seminal study.

The first term in Eq. (1) is derived formally in appendix A. It models the radiative forcing from aerosol–radiation interactions $\overline{F}_{\rm ari}$ as being proportional to $\overline{Q}_a$, following Charlson et al. (1992). The magnitude of $\overline{F}_{\rm ari}$ depends on a number of factors, such as the degree of cloud masking, the oxidation rate of SO_2, the composition of the aerosol, its covariability with meteorological conditions, and its lifetime. And although the individual factors can vary greatly, the atmosphere is effective at mixing in the phase space defined by these factors, so assuming that their net effect (i.e., the covariances among different terms) has varied little with time as compared to variations in $\overline{Q}_a$ is not as radical as it might seem. This justifies lumping them into a single time-invariant parameter α.

The second term in Eq. (1) is derived formally in appendix B. It models the radiative forcing from aerosol–cloud interactions $\overline{F}_{\rm aci}$. Physically one expects $\overline{F}_{\rm aci}$ to depend on the change in cloud droplet number concentrations N attributable to changes in the local aerosol burden B. A number of relationships between B and N have been proposed in the past, some of which are summarized by Storelvmo et al. (2009). Most adopt a power-law form, $N \propto B^x$. Assuming that the changing burden results solely from the changing source strength so that $\delta B \propto \delta Q$, it follows that $\delta N/N \propto x(\delta Q/Q)$. Algebraic increments in the source strength have diminishing returns; equivalently, the forcing from aerosol–cloud interactions changes arithmetically for geometric changes in the source strength. To capture these effects, it is proposed that the forcing $\overline{F}_{\rm aci}$ depends logarithmically on $\overline{Q}_a$ similarly to what has been assumed in other studies [cf. Fig. 6 in Boucher and Pham (2002) and Fig. 3 in Carslaw et al. (2013)]. Analogously to α, the parameter β subsumes a great many other processes, such as the covariance of cloud susceptibility and aerosol loading, which can be interpreted physically and whose net effect is assumed to have not varied on average, over the industrial period.

Despite its simplicity, Fig. 2 demonstrates that with a suitable choice for the free parameters, Eq. (1) provides a satisfactory model of the time history of $\overline{F}_{\rm aer}$. Through a different specification of its free parameters, Eq. (1) also provides a way to encapsulate uncertainty, as for instance would be represented by different reconstructions of $\overline{F}_{\rm aer}$ by different comprehensive models. A number of studies have emphasized that uncertainty in estimates of $\overline{F}_{\rm aci}$ arise from differences in the assumed strength of background aerosol burdens (e.g., Hoose et al. 2009; Carslaw et al. 2013). To represent this effect a natural source, $\overline{Q}_n$ has been introduced to parameterize the buffering effect of all natural aerosols in terms of an equivalent SO_2 source, which as such should be larger than estimates of the natural SO_2 source. It is important because it implies that uncertainty in the aerosol forcing is not linearly related to the central estimate of $\overline{F}_{\rm aer}$, as it would be if α and β were the only sources of uncertainty in the model. This behavior is in contrast to what is presented in the AR5 (IPCC 2013, see their Fig. SPM.5), which appears to be based on the assumption that the forcing uncertainty is linearly proportional to the forcing itself. The tendency of the uncertainty in the forcing to increase more rapidly than the forcing itself is evident in Fig. 2, where the spread in the forcing estimates (shown by the difference between the dashed lines) grows faster with time than the forcing itself, and is well established already by 1950.

3. Implications of the simple model for aerosol forcing

Considerable benefit can be derived by expressing $\overline{F}_{\rm aer}$ as a function of $\overline{Q}_a$. Emissions of SO_2 increased very rapidly from the early part of the twentieth century until about the mid-1970s when regulations began limiting

further emissions. Hence the emission history of SO_2 (the main precursor of anthropogenic aerosols) is very different from that of CO_2. SO_2 burdens have leveled off, or even fallen, with the introduction of measures to reduce pollution (Fig. 1), while CO_2 arising from anthropogenic activities continues to accumulate in the atmosphere. The short lifetime of aerosol particles and their precursors also means that the spatial patterns of forcing is disproportionately concentrated in the Northern Hemisphere for aerosol byproducts of combustion (and smelting), as compared to CO_2, which is long lived. Because of the disproportionate change in aerosol forcing to a unit perturbation in a pristine atmosphere, these differences would be expected to be amplified in the early part of the historical record, particularly in the Northern Hemisphere. For instance, in the period before 1950 the anthropogenic sulfate burden was increasing twice as rapidly as CO_2 when measured relative to their respective background burdens.

The period prior to 1950 is also interesting because there was marked warming in the early part of the century that appears difficult to reconcile with a very strong aerosol forcing. The median of 100 ensemble members from the HadCRUT4 dataset (Morice et al. 2012) suggest a 0.3-K warming. Indeed, it was this warming that motivated early speculation as to the role of rising concentrations of atmospheric carbon dioxide (Callendar 1938). As shown in Fig. 3, most of this warming occurred in a 30-yr period starting after the termination of a period of active volcanism and ending around 1950 when $\overline{Q}_a$ began to increase very rapidly. The temperature record also shows that the warming does not obviously originate in the Southern Hemisphere, as one might expect to happen if the global radiative forcing were positive, despite a negative forcing in the Northern Hemisphere. Combining Eq. (1) with estimates of forcing from long-lived greenhouse gases and chlorofluorocarbons (CFCs) suggests that for $\overline{F}_{\text{aer}} = -1.5\,\text{W m}^{-2}$ the net radiative forcing prior to about 1980 would have been negative (Fig. 4a). Given the history of observed warming, and after accounting for the volcanic activity in the three decades between the early 1960s and early 1990s, this does not seem plausible.

This line of argumentation can be developed to further bound the magnitude of an aerosol forcing whose historical evolution can be described by an equation of the form of Eq. (1). Supposing that, as is stated in the AR5 (IPCC 2013), it is extremely likely that *most* of the 0.5-K warming since 1950 can be attributed to anthropogenic activity, it seems equally unlikely that *none* of the 0.3-K warming between 1850 through 1950 can be attributed to anthropogenic forcing. Or put another way, it seems very unlikely that the natural contribution

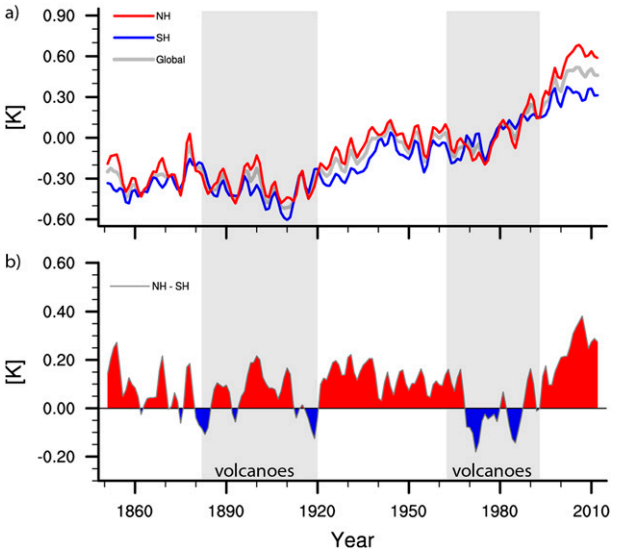

FIG. 3. (a) Time series of hemispherically averaged surface temperatures plotted as anomalies relative to the average over the 1961–90 period and (b) their difference. Data are from the HadCRUT4.2 dataset. Also shown are gray-shaded bands indicating periods of more active volcanism.

to the warming, generally thought to be due to a confluence of increased insolation during a quiescent period of volcanism (e.g., Suo et al. 2013), was so strong that it offset a negative anthropogenic forcing. This idea that the anthropogenic forcing was nonnegative at the end of the period of rapid warming in 1950 can then be used to provide tighter bounds on the present day magnitude of $\overline{F}_{\text{aer}}$. Choosing 1950 as an end year emphasizes (for the reasons discussed above) differences in aerosol versus greenhouse gas forcing. It also avoids the effects from very rapid increases in $\overline{Q}_a$ between 1950 and 1975 (choosing 1975 would yield a yet stronger constraint) that, convolved with the effects of volcanism starting with the large eruption of Agung in 1963, may have indeed produced a negative forcing.

To estimate a lower bound for the present-day $\overline{F}_{\text{aer}}$ from the constraint that the total forcing value in the year 1950 must be nonnegative, I evaluate Eq. (1) given $\overline{Q}_a$ in 1850, 1950, and 2005, with the parameters α, β, and $\overline{Q}_n$ chosen randomly from 10^5 draws of a prescribed distribution. I thereby simulate a wide range of possible relationships between $\overline{F}_{\text{aer}}$ and $\overline{Q}_a$ subject only to the form of Eq. (1). Values of α are chosen to vary so that their 2σ range for the present day forcing is between -0.1 and $-0.6\,\text{W m}^{-2}$. The value of $\overline{Q}_n$ is allowed to vary between 30 and 90 Tg $SO_2\,\text{yr}^{-1}$ (2σ), which is 50% larger than the values given by Carslaw et al. (2013) so as to account for nonsulfate sources of background CCN. Given a draw of α and $\overline{Q}_n$, the parameter β is chosen to sample present-day aerosol forcing ranging from

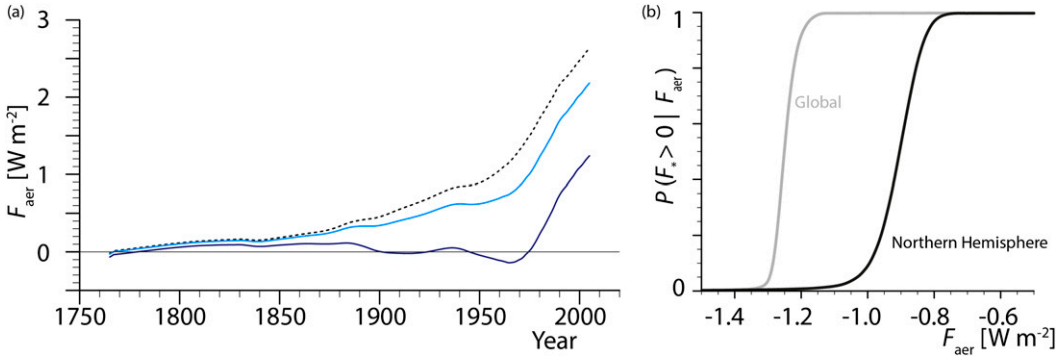

FIG. 4. (a) Net forcing resulting from long-lived greenhouse gases (including CFCs) alone (dashed), with $\overline{F}_{aer} = -0.5\,W\,m^{-2}$ (light blue), and with $\overline{F}_{aer} = -1.5\,W\,m^{-2}$ (dark blue). The weaker aerosol forcing estimate is based on Eq. (1) with the parameters of the central estimate in Fig. 2. The stronger aerosol forcing estimate adopts the parameters of the high forcing estimate in Fig. 2, with β increased so that the magnitude of $\overline{F}_{aer}$ maximizes at $1.5\,W\,m^{-2}$. (b) Probability that the net forcing over the globe (solid) or the Northern Hemisphere (dashed) between 1850 and 1950; F_*, is positive, given a random sampling of the model parameters conditioned on the global magnitude of $\overline{F}_{aer}$.

0 to $-1.5\,W\,m^{-2}$ and thus is also varied over a considerable range, whereby the upper and lower bounds in this range end up having no influence on the argument (i.e., one could sample much larger or much smaller forcing without changing the result).

The value of the aerosol forcing associated with an emission source equivalent to that in the year 2005 is equated with the present-day aerosol forcing, and denoted $\overline{F}_{aer}$. Estimates of the global aerosol forcing in 1850 and 1950 are combined with forcing from long-lived greenhouse gasses (CO_2, N_2O, CH_4, and CFCs) for those same periods to calculate the change in the net anthropogenic forcing. Other sources of anthropogenic forcing, such as ozone and land use, are not included here, but compensate one another to first order as discussed in appendix C. I denote the difference between the anthropogenic forcing by aerosols and long-lived greenhouse gases in 1950 and in 1850 by F_*. Based on the 10^5 different estimates of $\overline{F}_{aer}(\overline{Q}_a)$, I calculate the conditional probability, $p(F_* > 0|\overline{F}_{aer})$. This probability is presented graphically in Fig. 4b. It suggests that, to the extent that the aerosol forcing follows the form of Eq. (1), the requirement of a positive anthropogenic forcing in 1950 implies an aerosol forcing that is not more negative than $-1.3\,W\,m^{-2}$.

Idealized experiments using several comprehensive climate models suggest that the hemispheric temperature response is expected to follow the sign of the hemispheric forcing. Voigt et al. (2014a) explored the effect of asymmetric hemispheric forcing by perturbing the surface albedos differently in the different hemispheres (zero mean) in four different general circulation models running in an aquaplanet configuration, all with a very different representations of the tropical climate. The asymmetric forcing, ΔF in Table 1, is measured as the difference between the dark and bright hemispherically averaged radiative forcing and is calculated in a way that accounts for changes in cloudiness (Voigt et al. 2014b). The calculations demonstrate that a pronounced temperature difference results from an asymmetric forcing, suggesting that shifts in the Hadley cell (which is the prime way in which the atmosphere transports heat across the equator) does not completely compensate hemispheric forcing asymmetries. Based on this result, and the evidence that the warming in the early part of the century is, if anything, stronger in the Northern Hemisphere, I argue that the magnitude of $\overline{F}_{aer}$ can be further bounded by requiring the Northern Hemisphere value of F_* to be nonnegative.

The condition that F_* averaged over the Northern Hemisphere be nonnegative adds an additional constraint because the aerosol forcing is disproportionately concentrated in the Northern Hemisphere. To arrive at a lower bound on the forcing subject to this additional constraint, I model the Northern Hemisphere aerosol forcing as being proportional to its global value by a factor γ. Nine models [IPSL-CM5A-LR, CanESM2, CSIRO Mk3.6.0, Hadley Centre Global Environment Model, version 2–Atmosphere only (HadGEM2-A), GFDL CM3, MIROC5, FGOALS-s2, MRI-CGCM3, and BCC_CSM1.1] that as part of CMIP5 performed the SSTClim and SSTClimAerosol simulations are analyzed to estimate γ. For these models γ varies between 1.34 and 1.75, not including one model with a very small forcing and very large (7.8) ratio between the Northern Hemisphere and global forcing. Thus, to sample a wide range of uncertainty, I assume $\gamma = 1.5 \pm 0.4$ (2σ). A

TABLE 1. Summary of differences in surface temperatures between hemispheres asymmetrically forced by the indicated forcing ΔF. The sense of the temperature change follows that of the forcing. Calculations were performed for an aquaplanet with a mixed layer ocean whose depth is set to 30 m, and for which heat flux convergence is set to zero at each point. Experiments with a 50-m mixed layer ocean show a stronger response (5%–20%) but the forcing is unchanged. Model names follow the designation in Voigt et al. (2014a).

Model	ΔF (W m^{-2})	ΔT (K)
ECHAM6-TNT	1.4	0.78
ECHAM6-TTT	1.4	0.93
LMDZ5A	1.4	0.89
LMDZ5B	1.5	0.96

value of $\gamma = 1.5$ implies that the Northern Hemisphere aerosol forcing is 3 times as large as the Southern Hemisphere aerosol forcing. The strong hemispheric asymmetry in the forcing arises for reasons discussed earlier, and is consistent with patterns that robustly emerge from comprehensive modeling as discussed by Shindell et al. (2013). Further amplification in the hemispheric asymmetry of the forcing arises because extratropical forcing is more potent than tropical forcing, and any Southern Hemisphere aerosol forcing is concentrated in the broader tropics (Hansen et al. 1997; Kang and Xie 2014; Shindell 2014).

The conditional probabilities from this further constraint, shown by the dashed line in Fig. 4b, suggest that $\overline{F}_{\text{aer}}$ is unlikely to be below $-1.0\,\text{W m}^{-2}$. For $F_* > 0$ and $\overline{F}_{\text{aer}} < -1\,\text{W m}^{-2}$ very implausible parameter values are required, wherein almost all of the forcing is carried by the linear term in Eq. (1), that is, aerosol–radiation interactions.

In summary, a simple model of aerosol forcing [Eq. (1)], shown to be a good approximation of present-day understanding of aerosol processes, is used to revisit the lower bound on aerosol forcing. I use this model to interpret the time history of radiative forcing over the Northern Hemisphere prior to 1950. Based on this analysis I argue that an aerosol forcing less than $-1.0\,\text{W m}^{-2}$ is very unlikely. It seems likely that the 0.3-K rise of temperatures in the first half of the century likely has a naturally forced component, for instance from increasing insolation and the rebound from volcanic forcing. But a present-day aerosol forcing more negative than $-1.0\,\text{W m}^{-2}$ would imply that *none* of the rise in Northern Hemisphere surface temperatures during the 100-yr period from 1850 to 1950 could be attributed to anthropogenic forcing. This would imply a degree of natural variability that I find difficult to reconcile both with variability in comprehensive modeling (as discussed subsequently) and with the consensus that most of the post-1950 temperature rise can be attributed to anthropogenic causes.

4. Reconciling less negative aerosol forcing with physical understanding

In the AR5, the central estimate of $\overline{F}_{\text{aer}}$ was set, by expert judgement, to $-0.9\,\text{W m}^{-2}$, slightly less negative than what I posit for the lower bound. Below I review previous estimates of $\overline{F}_{\text{ari}}$ and $\overline{F}_{\text{aci}}$ using a more bottom-up approach in light of present-day observations. Based on this I argue that the AR5 best estimate is very much on the edge of what is plausible, as physical understanding constrained by present-day observations supports the revised lower bound such that $\overline{F}_{\text{aer}} > -1.0\,\text{W m}^{-2}$. I approach the problem in two parts, first by estimating the forcing from aerosol–radiation interactions, and second by estimating the forcing from aerosol–cloud interactions, as outlined below. I assume that the two contributions add linearly, although this will tend to overstate the forcing, as we know that stronger forcing from aerosol–cloud interactions implies brighter clouds, which implies (all else being equal) a smaller forcing from aerosol–radiation interactions.

a. *Aerosol–radiation interactions*

In Eq. (1) the model for $\overline{F}_{\text{ari}}$ is based on ideas introduced more than 20 years ago (Charlson et al. 1991, 1992). Following the approach of these authors, it can be shown (see appendix A) that α can be related to *effective* values of parameters whose *physical* value can be either measured or derived from first principles, whereby

$$\alpha \equiv \eta \left(\frac{C_r E_r K T_* \frac{3}{2} Y}{\Omega} \right). \tag{2}$$

In this expression Y is the effective sulfate yield from the oxidation of SO_2, with the factor of 3/2 accounting for the difference between the molecular weight of sulfate and SO_2; T_* is an effective lifetime, K an effective mass extinction, E_r the effective clear-sky fraction, Ω the surface area of the earth, and η the scaling of the $\overline{F}_{\text{ari}}$ from sulfate alone. One argument for comprehensive modeling approaches is that in characterizing the patterns of aerosol burdens, and their covariances with other fields, it provides a way to estimate the effective value of a given parameter given the physical value. In the original application of this approach important parameters were greatly overestimated (see the appendixes; see also Boucher and Anderson 1995), and the difference between the *physical* and *effective* value of a parameter were not properly

accounted for. As a result the estimate of the clear-sky forcing by Charlson et al. was about a factor of 5 larger than present-day estimates. Nonetheless, the form of Eq. (1) remains a powerful framework for interpreting aerosol forcing, as evidenced by its use to interpret results from comprehensive models (cf. Schulz et al. 2006; Myhre et al. 2013b).

The interpretive framework implied by Eq. (2) is routinely used to understand differences in more complex models and shows that models arrive at similar estimates of $\overline{F}_{\rm ari}$ from sulfate alone (about $-0.35\,{\rm W\,m^{-2}}$) in very different ways. Although some compensation among errors in the various components of $\overline{F}_{\rm ari}$ is expected on physical grounds (e.g., Boucher and Anderson 1995), the spread in different estimates of its constituent components undermines confidence in the apparent consensus emerging from comprehensive modeling (Schulz et al. 2006; Myhre et al. 2013b).

Adjusting the sulfate-aerosol forcing to account for other components of the aerosol, or from physical adjustments, amounts to estimating η. This is challenging, as it requires piecing together contributions taken from a very inhomogeneous sampling of models, with widely divergent estimates of individual forcing components (Shindell et al. 2013; Myhre et al. 2013b). For instance, estimates of the radiative forcing by nitrate vary by more than a factor of 10 in the most recent model intercomparison (Shindell et al. 2013). Nonetheless, taken at face values, most models estimate a positive contribution to $\overline{F}_{\rm aer}$ from nonsulfate aerosol[3] consistent with an all-aerosol $\overline{F}_{\rm ari}$ of about $-0.25\,{\rm W\,m^{-2}}$ (cf. Myhre et al. 2013a). A much more negative value ($-0.45\,{\rm W\,m^{-2}}$) of the all-aerosol $\overline{F}_{\rm ari}$ is given in the AR5 because that assessment gives more weight to strong nitrate forcing by a few models (Shindell et al. 2013) and includes a negative adjustment ($-0.1\,{\rm W\,m^{-2}}$) based on results from a single study, yet dismisses evidence of stronger positive adjustments from other studies (Boucher et al. 2013).

Observations of Earth's energy budget suggest that even the less negative, $-0.25\,{\rm W\,m^{-2}}$ estimate of $\overline{F}_{\rm ari}$ by comprehensive models may be too negative; as compared to observations of Earth's energy budget, the models reflect too much clear-sky radiation. Because the anthropogenic aerosol predominates in the Northern Hemisphere, the effect of aerosol forcing should be evident in the observed hemispheric asymmetry in the effective clear-sky albedo over the ocean A as a function of latitude φ. I denote this asymmetry by $A(\varphi)$. In computing $A(\varphi)$, only the clear-sky albedo is examined, so as to avoid complications from different cloud distributions, and only oceanic regions are compared, so as to minimize effects from differences in the underlying surface. The albedo itself is reconstructed from the annually averaged clear-sky irradiance provided by a 13-yr climatology derived from the satellite radiances measured by CERES and distributed in their edition 2.8 (Ed2.8) data collection (Loeb et al. 2009). By construction, for identical incident irradiance this effective albedo measures the annually averaged reflected radiation in clear skies over the ocean. Nonzero values of $A(\varphi)$ are interpreted as a measure of the difference in aerosol burdens between specific latitudes in the two hemispheres.

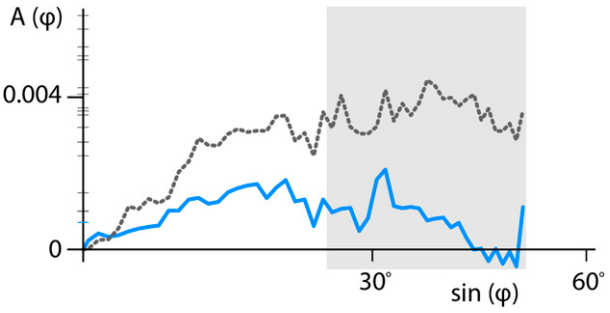

FIG. 5. Asymmetry $A(\varphi)$ in zonally and annually averaged albedo α over the oceans as a function of sine of latitude φ. CERES measurements (blue) and the multimodel mean (dotted), excluding one outlier model (MRI-CGCM3). The mean between 25° and 50° latitude of the individual models and of CERES is indicated by the minor tick marks on the ordinate axis.

For reasons elaborated on in the next section, it is assumed that, in regions where large asymmetries in the background natural aerosol are not expected, the asymmetry measures the anthropogenic aerosol burden. Consistent with this interpretation, Fig. 5 shows that the clear-sky albedo over the oceans in the Northern Hemisphere is, as expected, larger than that it is over the oceans in the Southern Hemisphere, although differences in the tropics likely reflect differences in the natural aerosol, for instance from mineral dust sources in North Africa. Averaging between 25° and 50° latitude, where anthropogenic sources are expected to be large but contributions from mineral dust should be less important, yields $\overline{A} = 0.77 \times 10^{-3}$ equivalent to a $-0.5\,{\rm W\,m^{-2}}$ difference in the reflected solar irradiance. CMIP5 models have a robustly larger value of $\overline{A}$ as compared to what is measured by CERES, roughly a factor of 5 for the AMIP simulations analyzed here (see Table 2). A smaller value of $A(\varphi)$ in the observations, as compared to the models, is consistent with the impression that the models overstate the downstream

[3] Zelinka et al. (2014) show that for the CMIP5 models performing aerosol only runs for forcing calculations, absorption makes $\overline{F}_{\rm ari}$ about 30% less negative than what one derives from scattering by the sulfate aerosol alone.

TABLE 2. AMIP simulations from the CMIP5 database (see also Table C1); tabulated are the analysis periods, the global surface and TOA albedo as well as the asymmetry of clear-sky ocean albedos ($\overline{A}$), and the radiative forcing associated with this asymmetry ($F^{\mathrm{clr}}_{\mathrm{asy}}$) for each of the models.

		Effective albedo			
Model name	Years	Surface	TOA	$\overline{A}$ (10^{-3})	$F^{\mathrm{clr}}_{\mathrm{asy}}$ (W m^{-2}; $25° \le \varphi < 50°$)
ACCESS1.0	1979–2008	0.146	0.177	2.82	−1.88
BNU-ESM	1979–2008	0.128	0.164	6.15	−4.03
CanAM4	1950–2009	0.155	0.176	4.25	−2.90
CCSM4	1979–2010	0.142	0.166	4.08	−2.78
CESM1(CAM5)	1979–2005	0.141	0.168	1.00	−0.58
CNRM-CM5	1979–2008	0.117	0.182	5.32	−3.75
CSIRO Mk3.6.0	1979–2009	0.149	0.188	5.79	−4.12
GFDL CM3	1979–2003	0.150	0.179	5.09	−3.34
GISS-E2-R	1951–2010	0.136	0.176	2.82	−1.94
HadGEM2-A	1978–2008	0.144	0.177	3.53	−2.35
INM CM4	1979–2008	0.183	0.188	1.01	−0.78
IPSL-CM5A-LR	1979–2009	0.152	0.171	3.72	−2.49
IPSL-CM5B-LR	1979–2008	0.151	0.171	3.64	−2.43
MIROC5	1979–2008	0.147	0.174	3.21	−2.15
MPI-ESM-MR	1979–2008	0.149	0.180	1.49	−0.85
MRI-CGCM3	1979–2010	0.175	0.190	2.46	−1.58
NorESM1-M	1979–2005	0.141	0.168	4.98	−3.48

influence of recent increases in Chinese SO_2 emissions. Over the past 10–15 years, a time period during which Chinese SO_2 emissions increased by 50%, satellite measurements show little evidence of a marked increase in aerosol optical depth or clear-sky reflectance over the open ocean downstream of East Asia (Shindell et al. 2013; Stevens and Schwartz 2012; Murphy 2013).

It is possible that biases in measurements of $\overline{A}$ arise for methodological reasons. For instance, systematic differences in the sea state between the hemispheres could cause $\overline{A}$ to depart from zero, even if the atmospheric composition were the same in both hemispheres. To address this possibility I also adopted the same procedure to compare the surface albedo asymmetry, which in the CERES data accounts for ocean color differences, and the wind speed dependence of surface albedo. Surface albedo asymmetries in the CERES (Ed2.7) surface radiation product are negligible, and show no evidence of the oceans in the Southern Hemisphere being brighter than their counterparts in the Northern Hemisphere in a way that would systematically bias the CERES estimates of $\overline{A}$ too low.

A possible bias associated with hemispheric biases in the identification of clear-sky scenes could also cause $\overline{A}$ to be underestimated by CERES. There are differences between the value of $\overline{A}$ derived from the CERES synoptic (SYN) data as compared to that derived from EBAF. The CERES SYN product uses a more conservative algorithm for identifying clear skies, as it requires the entire CERES (20-km nadir) footprint to be clear when deriving the clear-sky irradiances. To reduce the dependency on clear-sky fraction, the EBAF Ed2.8 product scales the irradiance data for the clear regions within CERES footprints from MODIS pixels identified as clear at 1-km spatial resolution. The SYN data have a twofold larger asymmetry, although it is still systematically smaller than that of the models. Inspection of the data shows the differences between the EBAF and SYN products to result from somewhat less reflected shortwave radiation poleward of 45°S over the Southern Ocean. In discussing this matter with the CERES team (W. Su 2014, personal communication) it was pointed out that near 45°S cloud fraction increases poleward along with the aerosol optical depth (as derived from MODIS). Hence a more conservative algorithm for identifying clear skies will underestimate aerosol optical depth, implying a darker Southern Hemisphere and a greater hemispheric asymmetry. These arguments lead me to believe that the EBAF data are more representative.

Related to this issue of cloud clearing, because the models define clear sky differently than the observations, the model asymmetry may be amplified by the clear skies being systematically more humid (cf. Sohn et al. 2010; Boucher and Quaas 2012) in the model, which would amplify (through a humidification effect) any background asymmetry. However, the relatively small (1 km) clear-sky footprint of the CERES EBAF data should mitigate against such effects. Additionally, the tendency for the EBAF product, which uses the less conservative (than the SYN product) cloud clearing algorithm, to be less asymmetric argues against the sampling playing a major role.

A further indication that something is amiss in the models rather than in the data arises from a simple

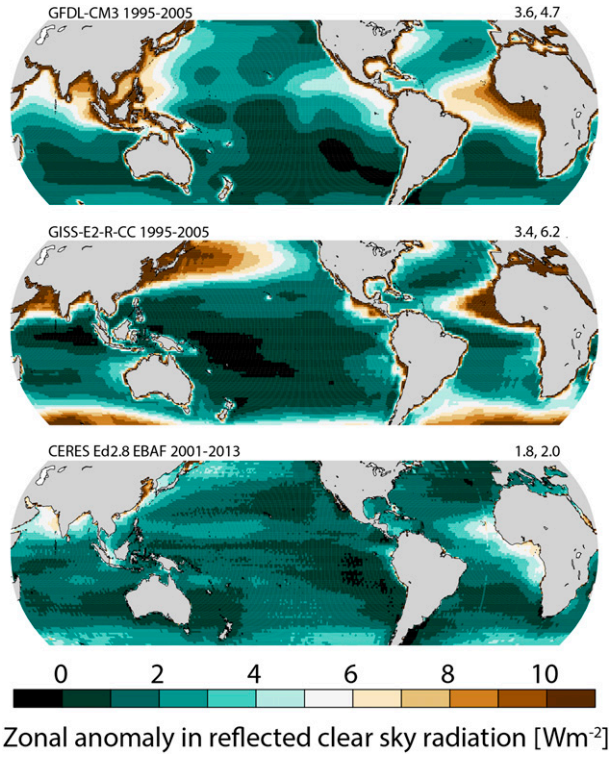

FIG. 6. Zonal anomaly in top-of-atmosphere clear-sky radiation from two models, (top) GFDL-CM3 and (middle) GISS-E-R-CC, and (bottom) as observed by CERES EBAF (Ed2.8). The anomalies are calculated with respect to a baseline as described in the text. The global and Northern Hemisphere average of the fields are given on the upper right of each panel.

inspection of maps of the outgoing clear-sky shortwave irradiances from the models and from CERES. In Fig. 6 the 10-yr average of irradiances from two models are compared with CERES. The two models were selected for presentation because they formed the basis of the assertion by Shindell et al. (2013) that changes in the pattern of aerosol emissions have caused $\overline{F}_{aer}$ to increase in magnitude since 1980. Compared to the average CMIP5 model these models have a relatively advanced representation of aerosol processes, and a reasonably good (as compared to data) representation of the pattern and magnitude of clear-sky aerosol radiative effects relative to other models. To mitigate against the influence of surface albedo biases the clear-sky aerosol radiative effects are equated with anomalies from a latitudinally varying (but hemispherically symmetric) baseline value. The baseline is calculated as the average of the vigintile of the least reflective points at a given absolute latitude (for the period between 1860 and 1870 in the models). Using the absolute latitude ensures that hemispheric asymmetries are not artificially introduced when constructing the anomalies and implies that at most 5% of the points will have a negative value. Figure 6 shows that the historical simulations by both models clearly overstate the amount of outgoing shortwave irradiance in the Northern Hemisphere. In GFDL CM3 the differences with the CERES data are pronounced in the tropical oceans south of Asia and west of Africa. In the GISS model there is a more (as compared to CERES) pronounced anomaly in reflected shortwave radiation over the North Pacific Ocean, climatologically downwind of Asia. But even in regions such as the North Atlantic both models show the signature of a more reflecting aerosol plume climatologically downwind of North America. The models are as different from one another as they are from CERES, further supporting my contention that more is amiss with the models than the data. The models, however, also show consistent differences with CERES, in that signatures of aerosols are much more localized near the continents in the data. The data thus paint a picture that is consistent with previous analyses, which have explored whether or not maritime areas downwind of regions that have experienced a large increase in aerosol and aerosol precursor emissions have also experienced large changes in aerosol optical depth (Shindell et al. 2013; Stevens and Schwartz 2012; Murphy 2013). These studies show little evidence of pronounced trends away from the source regions, suggesting that the models that do show such trends are overstating the effect of anthropogenic emissions on clear-sky radiances.

The CERES data can be used to construct a rough estimate of $\overline{F}_{ari}$ by scaling the observed value of $\overline{A}$ by the ratio of $\overline{F}_{ari}$ to $\overline{A}$ from the modeling estimates. For the Max Planck Institute for Meteorology (MPI-M) Aerosol Climatology (MAC-v1.0; Kinne et al. 2013), which is based on present-day observations taken from AERONET and is one of the less biased estimates of the clear-sky aerosol effect (see the values for the MPI-ESM, which uses this climatology, in Table 2), this ratio is $0.477\,W\,m^{-2}$, which suggests a value of $\overline{F}_{ari} = \overline{A}(\overline{F}_{ari}/\overline{A})_{MAC} = -0.15\,W\,m^{-2}$. To arrive at this number I assume an effective clear-sky fraction of 0.65, somewhat larger than the mean (0.62) inferred from a recent comparison of three-dimensional radiative transfer calculations using the observed climatology of clouds (Stier et al. 2013). Assuming a smaller clear-sky fraction would yield an even less negative estimate of $\overline{F}_{ari}$.

In summary, the data provide no evidence that the models produce an insufficiently negative estimate of $\overline{F}_{ari}$, and considerable evidence that they may be substantially overstating the magnitude of $\overline{F}_{ari}$. Based on this I believe that a value of $\overline{F}_{ari}$ is very likely to be less, rather than more, negative than the value ($-0.25\,W\,m^{-2}$) taken from the modeling, so that $\overline{F}_{ari} > -0.25\,W\,m^{-2}$ would appear to be a reasonable upper bound on the radiative forcing from aerosol–radiation interactions alone.

b. Aerosol–cloud interactions

In Eq. (1) the term representing $\overline{F}_{\rm aci}$ can likewise be related to physical models of aerosol–cloud interactions. To do so I must assume that changes in cloud macrostructure (lifetime effects) that accompany anthropogenic perturbations to the cloud-active aerosol either are small or scale with changes in cloud microstructure. Given a poor understanding of the controls on cloud amount, and the lack of empirical evidence for cloud macroscopic changes as a function of aerosol concentrations, this is not an unreasonable assumption. Skeie et al. (2011) and Hansen (2005) have used similar approaches.

In this case, to estimate the average forcing $F_{\rm aci}$ it is sufficient to estimate the change in the local shortwave cloud radiative effect R that is attributable to a perturbation in the sulfate aerosol. Here again I follow the approach first outlined by Charlson et al. (1992). For an overcast layer $R = E\ln N$, where E denotes an efficiency, which for a given cloud type can be derived from radiative transfer calculations, and N denotes the cloud droplet concentration. Assuming that an aerosol perturbation only affects N, then

$$F_{\rm aci} = I_{\rm cld}\delta R = -I_{\rm cld}E\left(\frac{\delta N}{N}\right), \tag{3}$$

where $I_{\rm cld}$ is a weighting function. If changes in R as a function of N were the same for all clouds, $I_{\rm cld}$ would simply be zero or one depending on whether or not cloud is present. Cloud macrophysical changes, which effect E, can be accounted for by allowing $I_{\rm cld}$ to adopt values other than zero or one.

It is customary to think of cloud–aerosol interactions in terms of their effect on stratiform cloud layers, such as maritime stratocumulus, or stratiform cloud regions associated with shallow convection in the tropics, or postfrontal regions in the extratropics. There is good reason for this, as the relatively low optical thickness of these clouds, and the pristine environment in which they are found, make them particularly susceptible to perturbations in their droplet numbers, as evidenced by ship tracks. For a typical subtropical stratocumulus layer, analytic arguments (Charlson et al. 1992) and radiative transfer modeling can be used to estimate $E = 22\,{\rm W\,m^{-2}}$ (see appendix B). To estimate $\overline{F}_{\rm aci}$ one must average Eq. (3) over the globe and over time, which results in

$$\overline{F}_{\rm aci} = -CE\left(\frac{\delta\overline{N}}{\overline{N}}\right), \tag{4}$$

where an effective cloud fraction,

TABLE 3. Sulfate radiative forcing efficiency, $CE = \overline{F}_{\rm aci}(\overline{N}/\delta\overline{N})$, taken from the calculations tabulated in Table 1 of Storelvmo et al. (2009). To calculate C a value of $22\,{\rm W\,m^{-2}}$ is assumed for E. The nomenclature for the origin for each estimate follows that of Storelvmo et al.

Origin	$\overline{F}_{\rm aci}$	$\delta\overline{N}/\overline{N}$	CE	C
BL95	−0.91	0.197	−4.61	0.21
J01	−0.97	0.369	−2.63	0.12
M02	−1.94	0.776	−2.47	0.11
DO5	−0.62	0.251	−2.50	0.11

$$C \approx \overline{I}_{\rm cld}\left(1 - \frac{\overline{N'(\delta N)'}}{\overline{N}\,\delta\overline{N}} + \frac{\overline{I'_{\rm cld}(\delta N)'}}{\overline{I}_{\rm cld}\delta\overline{N}} - \frac{\overline{I'_{\rm cld}N'}}{\overline{I}_{\rm cld}\overline{N}}\right), \tag{5}$$

can be derived (see appendix C) by expanding the spatially and temporally varying terms in Eq. (3) into a mean and fluctuating component. It thus accounts for covariances that arise as a result of the averaging. These covariances, which were neglected by Charlson et al. (1992), can be considerable and generally act to make C less than $\overline{I}_{\rm cld}$ (i.e., cloud droplet perturbations are likely to be larger in arid regions where $I'_{\rm cld} < 0$ and $N' > 0$). The model of $\overline{F}_{\rm aci}$ used in Eq. (1) follows directly from Eq. (4) with $\delta\overline{N}/\overline{N} \propto \ln(\overline{Q}_a/\overline{Q}_n + 1)$. Because most of the covariances contributing to C are expected to be negative, the effective cloud fraction C will be smaller than the actual cloud fraction, which (assuming only low, liquid clouds are susceptible to aerosol perturbations and following calculations presented in the appendixes) is about 0.4.

The factor CE can be inferred from the literature. Storelvmo et al. (2009) used the ECMWF Integrated Forecasting System, which compared to many climate models has a relatively good representation of clouds, to explore the effect of different parameterizations of the cloud droplet concentrations on $\overline{F}_{\rm aci}$, given prescribed monthly concentrations of the aerosol. They did not calculate CE but it can be inferred from their Table 1, which along with their original calculations is reproduced in Table 3 herein. Three of the parameterizations produce very different predictions of the baseline droplet concentration, and a factor of 3 difference in $\delta\overline{N}/\overline{N}$, yet very consistent estimates of the effective cloud fraction, with $CE = 2.53 \pm 0.09\,{\rm W\,m^{-2}}$ or $C = 0.12$. The difference between these parameterizations and the one (which they call BL095) with $C = 0.21$ is an apparently much larger sensitivity to small changes in the sulfate mass concentration in the outlier (BL95) parameterization (e.g., Fig. 1 in Storelvmo et al. 2009), so that $\overline{F}_{\rm aci}$ will receive contributions from relatively remote regions. These results suggest that spatial heterogeneities act on

their own to reduce the effective cloud fraction by a factor of 4 (from 0.4 to 0.1). Accounting for temporal variability on submonthly time scales would lead to a further reduction in the effective cloud fraction,[4] so that even after accounting for the roughness of this analysis it seems hard to imagine values of C more than 50% greater than one finds after accounting only for spatial heterogeneities (i.e., $C < 0.15$). These results also suggest that to improve estimates of CE more attention should be paid to the effects of temporal variability [which Storelvmo et al. and Carslaw et al. (2013) and many other studies neglect] as well as the degree to which small changes in the aerosol loadings in remote regions project on cloud droplet changes.

Inferences from an analysis of observed cloud-radiative effects provide further support for $C \approx 0.1$. For a typical stratocumulus cloud with an insolation weighted solar zenith angle of 43.66°, radiative transfer calculations yield $R = -115\,\mathrm{W\,m^{-2}}$ for $N = 100\,\mathrm{cm^{-3}}$, or $R = -100\,\mathrm{W\,m^{-2}}$ for $N = 50\,\mathrm{cm^{-3}}$—a large value. As a reference, the globally averaged shortwave cloud radiative effect from CERES is $-46\,\mathrm{W\,m^{-2}}$. Hence if one assumes that such stratiform clouds contributed a cloud fraction of 0.3 they would be responsible for roughly 75% of the globally averaged forcing. Given the very large shortwave cloud radiative effect arising from much deeper and ice containing clouds, such a large contribution to the global cloud radiative effect is unreasonable (i.e., the low cloud fraction should be substantially less than 0.3). An effective stratiform cloud fraction of 0.3 also appears large when one considers that the average cloud-radiative effect over the subsiding regions of the tropical oceans is closer to $-20\,\mathrm{W\,m^{-2}}$ so that the equivalent stratocumulus cloud fraction in subsiding regions alone would be about 0.2. Given that subsiding air covers about 60% of the globe, and in regions of upward motion high clouds will increasingly mask changes to low cloud radiative effects, this implies an equivalent stratocumulus cloud fraction of 0.12, similar to Storelvmo et al. (2009). Even limiting oneself to a consideration of the climatological stratocumulus regions, defined as subsiding regions where the lower tropospheric stability is larger than 18 K (e.g., Klein and Hartmann 1993; Medeiros and Stevens 2011; Medeiros et al. 2015), and which cover about 30% of Earth's ocean, the cloud radiative effect is still $-45\,\mathrm{W\,m^{-2}}$. This implies an effective cloud fraction of 0.5 over these stratocumulus regions or about 0.15 overall. Using a different approach Wood (2012) arrives at a similar value. Because one does not expect every stratocumulus cloud on Earth to experience a change in its droplet concentrations on the order of the mean global change (which is mostly concentrated over land in the Northern Hemisphere) I find it difficult to make the case for a value of $C > 0.1$ and believe that $C = 0.15$ is a reasonable upper bound.

Estimates of C from comprehensive modeling, inferences from aerosol climatologies, and observations of cloud-radiative forcing are thus surprisingly consistent, indicating $C < 0.1$, more than a factor of 3 smaller than what was assumed by Charlson et al. (1992). In addition to suggesting that early estimates of $\overline{F}_{\mathrm{ari}}$ were too large, the consistency of the different estimates of C implies that most of the differences among models arises from differences in their prediction of $\delta\overline{N}/\overline{N}$ or from macroscopic changes in cloudiness (lifetime effects) that act to increase C. Changes in the average macroscopic properties of cloud fields as a result of an aerosol perturbation have been introduced in some comprehensive models, in a way that effectively increases C, but robust evidence for such effects is lacking (Stevens and Feingold 2009; Boucher et al. 2013).

To estimate $\overline{F}_{\mathrm{aci}}$, it is thus necessary to know $\delta\overline{N}/\overline{N}$. Based on hemispheric differences in measurements of non-sea-salt sulfate in remote locations Charlson et al. (1992) estimated $\delta\overline{N}/\overline{N} = 0.15$. Global modeling does not provide particularly robust estimates of cloud-droplet number concentrations, partly because these quantities may be tuned to help ensure that the radiation budget at the top of the atmosphere matches the observations even if the clouds do not (Nam et al. 2012), but more fundamentally because the processes that control cloud droplet number are not well represented by global models. So it is not surprising that even in the most ''advanced'' comprehensive models participating in the Aerosol Comparisons between Observations and Models (AEROCOM) project, $\overline{N}$ varies by more than a factor of 6 (from 19 to $122\,\mathrm{cm^{-3}}$) and $\delta\overline{N}/\overline{N}$ varies by more than a factor of 30 (from 0.06 and 2.25). Likewise, for a recent intercomparison with field measurements in the spatially extensive stratocumulus decks of the southeastern Pacific Ocean, predictions of droplet concentrations by state-of-the-art models varied by more than an order of magnitude and systematically underestimated the observed concentrations (Wyant et al. 2015). Even using a single model with the same aerosol perturbation (Storelvmo et al. 2009) showed that the parameterization of aerosol–cloud interactions alone can produce a fourfold difference in $\delta\overline{N}/\overline{N}$. Put bluntly, there is no evidence that the quantitative dependence of cloud droplet numbers on aerosol and aerosol precursor emissions is reliably represented by comprehensive modeling.

[4] Some modeling groups (e.g., Déandreis et al. 2012) are beginning to explore the role of temporal covariances in their models. In this respect critical comparisons between the modeling and data would help increase confidence that modeled signals are capturing something fundamental.

Some sense of the susceptibility of droplet concentrations to large-aerosol perturbations in pristine environments is provided by ship-track data. Retrievals of droplet sizes by satellite show that in detectable ship plumes the effective radius is reduced by 20% on average, equivalently $\delta\overline{N}/\overline{N} = 0.6$ (Christensen and Stephens 2011). In situ measurements are consistent with these satellite-derived estimates (Chen et al. 2012). Because ship tracks are favored in pristine environments and represent a very intense *local* perturbation, the value of $\delta\overline{N}/\overline{N} = 0.15$, which was originally suggested by Charlson et al. (1992) based on hemispheric measurements of non-sea-salt sulfate, does not seem too small.

Even allowing for what I believe to be an unrealistically large (factor of 2) uncertainty in my estimate of $\delta\overline{N}/\overline{N}$ implies that $\overline{F}_{\rm aci} > -0.75\,{\rm W\,m^{-2}}$, where this lower bound is derived by assuming that uncertainty in the estimate of C is independent of uncertainty in my estimates of $\delta\overline{N}/\overline{N}$. That my estimate of a lower bound for $\overline{F}_{\rm aci}$ is less negative than the central estimate of Charlson et al. (1992), who adopted an identical approach and who employed the same value of $\delta\overline{N}/\overline{N}$, is attributable to those authors' adoption of an unrealistically large value of C, which in part stems from their failure to account for covariances between aerosol perturbations and cloud incidence. Larger estimates of $\overline{F}_{\rm aci}$ from comprehensive modeling are often reported (e.g., Carslaw et al. 2013), but I believe that this reflects an unrealistic sensitivity of $\overline{N}$ to aerosol and aerosol precursor emissions in the comprehensive modeling and/or a failure to account for submonthly covariability between aerosols and their environment.

c. *New bounds on aerosol forcing*

Taken together my revised and observationally constrained estimates of $\overline{F}_{\rm ari}$ and $\overline{F}_{\rm aci}$ support the lower bound of $-1\,{\rm W\,m^{-2}}$, which was derived in the previous section based on a consideration of the historical record of globally averaged surface temperatures. When this lower bound is combined with bounds on aerosol forcing derived from observations of Earth's energy budget since 1950 (Murphy et al. 2009), my analysis suggests that

$$-1.0 < \overline{F}_{\rm aer} < -0.3\,{\rm W\,m^{-2}}. \qquad (6)$$

This represents a nearly threefold reduction in the uncertainty in aerosol forcing as compared to that given in the IPCC Fifth Assessment Report.

5. Reconciling less negative aerosol forcing with comprehensive modeling

The above analysis begs the question as to why comprehensive models are able to reasonably simulate the twentieth-century trends in globally averaged surface temperatures, despite values of $\overline{F}_{\rm aer}$ more negative than the lower bound postulated above. Zelinka et al. (2014) diagnose $\overline{F}_{\rm aer} = -1.4 \pm 0.56\,{\rm W\,m^{-2}}$ for a subset of nine CMIP5 models that perform idealized aerosol forcing experiments, and Wilcox et al. (2013) show that models that include aerosol–cloud interactions better represent interdecadal variability in the historical mean surface temperatures. Likewise, Ekman (2014) also shows that models with more advanced representations of aerosol–cloud interactions better capture the observed distribution of latitudinal temperature trends between 1965 and 2004. Using the entirety of the observed record can be misleading because during the latter part of the record the forcing from long-lived greenhouse gases dominates, so that if the models are too sensitive to greenhouse gas forcing a more negative aerosol forcing will lead to a better match between observed and simulated temperatures. Moreover, large differences in the treatment of volcanic forcing are difficult to disentangle from other the effects of anthropogenic forcing.

My arguments would suggest that to judge whether the simulated values of $\overline{F}_{\rm aer}$ are too negative, it would be more insightful to look at the models during a period with relatively little volcanism, and when aerosol forcing was more commensurate with greenhouse forcing in magnitude. The years of rapid warming between 1920 and 1950 define just such a period. Figure 7 suggests that the models warm too little during this period, consistent with an aerosol forcing that is too negative. The figure presents globally averaged surface temperatures from the first ensemble member of 35 model configurations that submitted a historical simulation to the CMIP5 archive. For each simulation a decadal temperature anomaly with respect to that simulation's 1961–90 mean temperature is calculated, and the distribution of decadal temperatures is plotted along with the median global temperature anomaly (with respect to the same period) from the median of a 100-member ensemble taken from the HadCRUT4 dataset. Overall the models provide a plausible representation of the instrumental record as a whole. However, there is an indication that the models systematically underestimate the warming in the 30-yr period between 1920 and 1950.

The slower warming during this period may be a consequence of internal variability, so that the average model warms less than what is observed. But the reduced warming simulated during this period is not simply a property of the multimodel mean. Only three or four of the 35 models (Fig. 7b) simulate as much warming as is observed during the period between 1920 and 1950, while six models show essentially no warming (or even cooling) during this 30-yr period. This result is consistent with the hypothesis that the aerosol forcing in the models is too

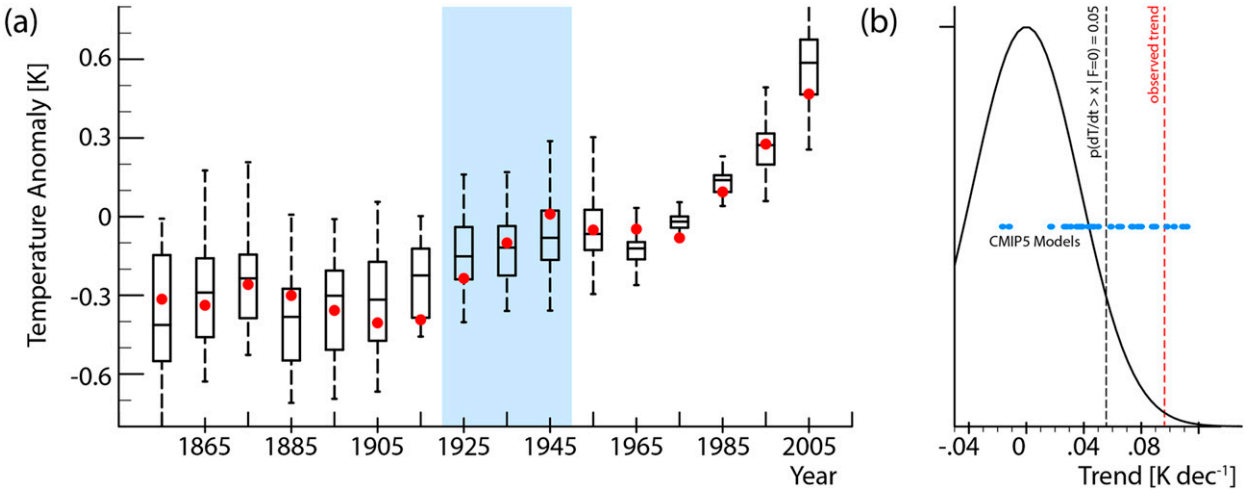

FIG. 7. (a) Decadal temperature anomalies with respect to the 1961–90 period for the CMIP5 historical simulations (box and whiskers) and from the HadCRUT4 dataset (red dots, median estimates). (b) Decadal trends (blue dots) for individual CMIP5 models during the 30-yr period 1920–50 overlain on the normal distribution of possible trends (with a standard deviation of 0.037 K decade^{-1} taken to be equal to the standard deviation of the regressed 1920–50 trends in a 100-member CMIP5 historical ensemble) for the case of no forcing. The 0.95 value of the cumulative distribution is shown by the black dashed line and the observed trend by the red dashed line. A naturally forced trend larger than the distance between the black line and the red line would make it impossible to rule out the possibility (at the 5% level) that the observed trend arose independently of anthropogenic forcing.

negative. Nonetheless, to explore the role of natural variability more thoroughly I analyze a 100-member ensemble of the latest version of the MPI Earth System Model (MPI-ESM, version 1.1), which was run for the period between 1850 and 2005. The 30-yr trends from this ensemble have been calculated by regressing annually averaged global surface temperatures against time for the period between 1920 and 1950. The standard deviation of the regressed trends is 0.037 K decade^{-1}. Assuming normally distributed trends, if the net forcing were negative the probability that the trend would be as large as observed (0.095 K decade^{-1}) is less than 0.5% (2.6σ). Some of the warming prior to 1950 is likely to have a naturally forced component, as insolation is believed to have increased during this period (Suo et al. 2013). The 100-member MPI-ESM historical ensemble thus suggests that the naturally forced trend would have to have accounted for about half of the observed trend for it to be explainable at the 10% level without any contribution from anthropogenic forcing.

Based on these measures of variability and noting that not all of the CMIP5 models are likely to have excessive aerosol forcing, my analysis supports the argument that the aerosol forcing in the CMIP5 ensemble is too negative. It would be interesting to more systematically test these ideas by running large ensembles in more models, ideally with the same (or at least a well characterized) aerosol forcing. Another way to develop this line of argumentation further would be to contrast the temperature trends in the present period, for which there have also been no major volcanoes since the eruption of Mt. Pinatubo in 1991 and aerosol forcing has been relatively constant, with those between 1920 and 1950. Here again, however, comparisons among models require a good characterization of the aerosol forcing applied to the models for both periods.

As alluded to in the last paragraph, one difficulty with interpreting the CMIP models is their very different representations of non–greenhouse gas forcing, particularly associated with aerosols. Interpreting the simulations in terms of the observational record thus convolves differences in how individual model configurations are forced with differences in their climate response. Except for the small subset of the models that performed dedicated experiments (cf. Zelinka et al. 2014) designed to assess the aerosol forcing in models, it is not possible to diagnose differences in $\overline{F}_{\text{aer}}$ across the models. And even for this subset, only the present-day forcing is calculated.

To circumvent this problem I estimate differences in the aerosol forcing across the CMIP5 ensemble by calculating the anthropogenic contribution to the asymmetry parameter, denoted A_a, as a function of time for the CMIP5 historical simulations. This is possible because for the historical simulations I can remove the contribution of the natural aerosol to A by subtracting the background asymmetry (not just the background clear sky) from a period late in the nineteenth century (1860–70) when aerosol forcing and volcanic activity were believed to be small. The results of this calculation are presented in Fig. 8. As a comparison, models submitting historical

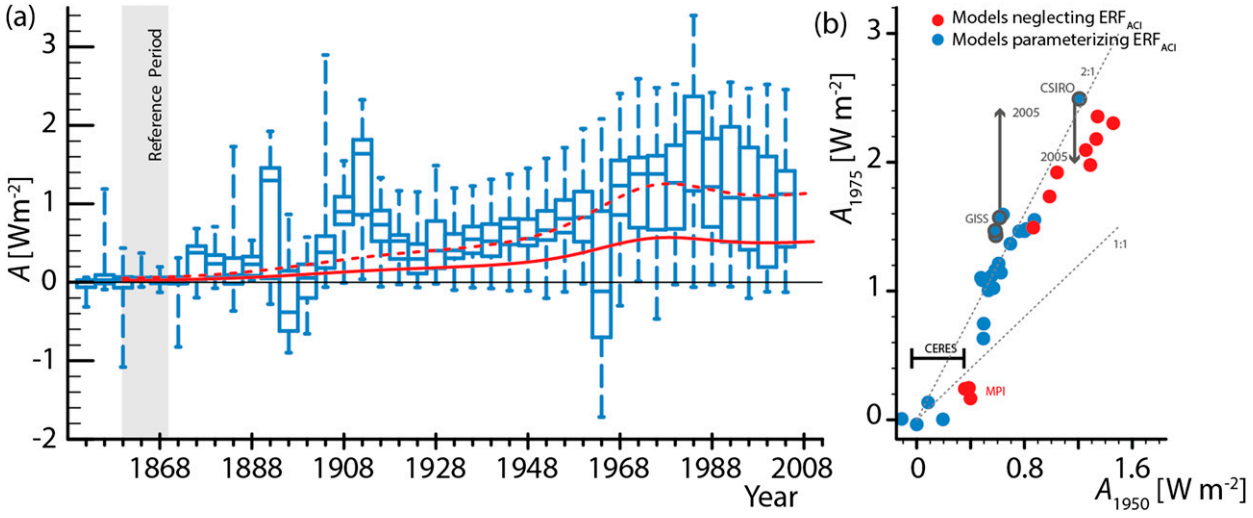

FIG. 8. (a) Clear-sky ocean asymmetry A calculated from CMIP5 historical simulations for 4-yr intervals after removing the asymmetry from the background aerosol estimated as the mean for the period between 1850 and 1870. The solid red line shows the asymmetry parameter (including natural background aerosol) from CERES scaled backward in times using changes in emissions of SO_2 relative to the present day. The dashed red line is the same as the solid, but multiplied by a factor of 2.2. (b) Clear-sky asymmetry parameter A for CMIP5 models (with the effects of natural aerosols removed) for the 10-yr period centered around 1975 vs the value for the 10-yr period centered around 1950. Models without aerosol–cloud interactions are denoted by red dots; models that include aerosol–cloud interactions are shown by blue dots. Also shown are estimates of A from CERES (without removing natural aerosols) for the present day. Because SO_2 emissions double between 1950 and 1975 if the asymmetry were to follow the anthropogenic emissions of SO_2 the models would be expected to scatter along the 2:1 line.

simulations with only natural forcing (historicalNat simulations) are also evaluated by calculating the change in total reflected clear-sky radiation over oceans (Fig. 9a) and the hemispheric asymmetry A_a (Fig. 9b). In these historicalNat simulations periods of volcanic activity are readily evident (Fig. 9a) but outside of these periods A_a is nearly zero, with little evidence of a trend, as one would expect as by definition A_a should be zero. This analysis thus supports the idea that A_a in the historical forcing experiments indeed measures the anthropogenic aerosol forcing. Figure 8 further shows that in the historical simulations A_a scales well with $\overline{Q}_a$, but it is more than twice as large as is observed. This discrepancy is even more striking when it is recalled that in the CERES data the contribution of natural aerosols to A has not been removed.

To more quantitatively compare the scaling of A_a with $\overline{Q}_a$, the value of A_a for the decade centered around 1975 is compared with the value of A_a for the decade centered around 1950. During this 25-yr period $\overline{Q}_a$ roughly doubled (Fig. 1), as does A_a in most of the models (Fig. 8b). Because $\overline{F}_{\rm aer}$ is expected to scale with changes in clear-sky radiation, this finding supports one of the premises of my argument, namely that $\overline{F}_{\rm aer}$ scales well with $\overline{Q}_a$. Figure 8 also illustrates the very large differences in apparent clear-sky aerosol forcing among the models. The effects of these differences on the total aerosol forcing are likely ameliorated by the tendency of models without a representation of aerosol–cloud interactions to have larger values of A_a and hence larger values of $\overline{F}_{\rm ari}$. Because in the absence of aerosol–cloud interactions $\overline{F}_{\rm aer}$ scales linearly with $\overline{Q}_a$, the magnitude of the aerosol forcing in these models is not as strongly constrained by the arguments of section 3 as are those models that include a representation of aerosol–cloud interactions.

6. Findings and implications

A simple model of aerosol forcing, shown to be consistent with present-day understanding of aerosol processes, is used to revisit the lower bound on aerosol forcing. I use this model to interpret the time history of radiative forcing over the Northern Hemisphere prior to 1950. Based on this analysis I argue that an aerosol forcing less than $-1.0\,{\rm W\,m^{-2}}$ is very unlikely. A more negative aerosol forcing would imply that none of the roughly 0.3-K rise in Northern Hemisphere surface temperatures during the 100-yr period from 1850 to 1950 could be attributed to anthropogenic forcing, which seems implausible. This lower bound is shown to be consistent with bottom-up estimates derived from physical understanding of aerosols and constraints from observations of Earth's energy budget, the amplitude of cloud droplet concentration changes associated with strong local forcing (ship tracks), and patterns of aerosol perturbations taken from comprehensive modeling. The

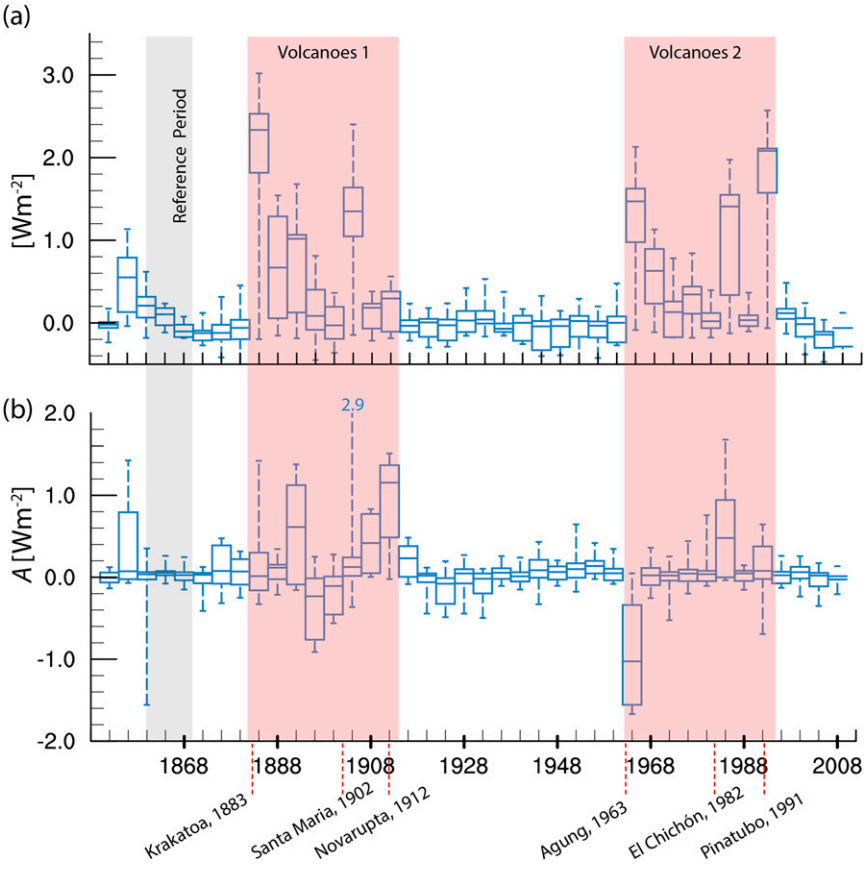

FIG. 9. (a) Change in globally averaged reflected clear-sky shortwave radiation over the ocean relative to a reference period (1860–70) for CMIP5 historical simulations with natural forcing only (historicalNat). (b) Change in asymmetry A relative to the reference period for CMIP5 historical simulations with natural forcing only.

argument of a weaker (less negative) aerosol forcing is also consistent with the tendency of comprehensive modeling to underestimate the warming in the period between 1920 and 1950, even after accounting for natural variability. In conclusion: three different lines of evidence provide support for an aerosol forcing less negative than $-1.0\,\mathrm{W\,m^{-2}}$. If one adopts an upper bound for the aerosol forcing of $-0.3\,\mathrm{W\,m^{-2}}$, based on an analysis of Earth's energy budget since 1950, this suggests that the radiative forcing from the anthropogenic aerosol is very likely (90%) to be between -0.3 and $-1.0\,\mathrm{W\,m^{-2}}$.

This range for the present-day aerosol forcing is consistent with, but considerably narrower than, the estimate of that same forcing in the IPCC AR5. The central estimate from the AR5 ($-0.9\,\mathrm{W\,m^{-2}}$) is also consistent with the present forcing range. Nonetheless, the arguments I adopt based on an analysis of the forcing prior to 1950 raises the question as to how the AR5 (e.g., Fig. 8.18 and Annex II therein) can so comfortably accommodate a very negative aerosol forcing without the appearance of a negative trend in the net anthropogenic forcing over the first 100 years (1850–1950) of the historical period. There are two explanations for this apparent inconsistency. The first is that my less negative forcing arises from my assertion that the Northern Hemisphere forcing must be positive between 1850 and 1950, and the AR5 shows global forcing. An argument that only considers the global forcing yields a somewhat more negative bound of $-1.3\,\mathrm{W\,m^{-2}}$. The second explanation is that the time series of aerosol forcing provided in the AR5 is unusual, and I believe unrealistic. The AR5 forcing is estimated to have increased as much between 1750 and 1850 as it did between 1850 and 1940 despite the fact that anthropogenic emissions of SO_2 increased elevenfold as much in the latter period as compared to the earlier period. One might be tempted to interpret this an extreme example of the nonlinearity in the forcing response to emissions expected from aerosol–cloud interactions, but this seems difficult to reconcile with a 30% increase in the forcing efficiency after 1940, even well before the pattern of global emissions began changing substantially.

At this point it seems worthwhile to step back and adopt a different perspective, as the present work raises the question as to why we, in the first place, think that aerosol forcing might be more negative than about $-1\,\mathrm{W\,m^{-2}}$. Just because we cannot model solar irradiance from first principles accurately is not a good basis for assuming that solar forcing before 1950 is hugely uncertain, so why should such an argument apply to estimates of aerosol forcing? Forcing estimates based on simple physical reasoning (Charlson et al. 1992) once motivated the consideration of a large and negative aerosol forcing, but these arguments are now shown to actually be consistent with forcing of a much smaller magnitude. Comprehensive modeling readily produces very negative estimates of aerosol forcing, but its quantitative representation of the distribution of important aerosol properties is not credible (e.g., Figs. 5 and 6) and is dependent on ever more speculative effects that are increasingly contradicted by finescale modeling (Stevens and Feingold 2009). In the present work it is shown that the models produce an anthropogenic aerosol signal that is distributed much more broadly over the World Ocean than is observed (e.g., Fig. 9.28 in Flato et al. 2013) and poorly represent what little we know about present-day droplet concentrations. Moreover, because $\overline{F}_{\mathrm{aci}}$ is expected to depend logarithmically on local perturbations to droplet concentrations, capturing the covariance between aerosols and clouds, not just spatially but temporally, is crucial to estimates of the forcing, and in this there is little basis for trusting estimates from comprehensive modeling. Even for relatively straightforward quantities where models appear to agree, such as the estimate of the forcing from sulfate aerosol–radiation interactions, agreement in a final number belies a divergence of estimates in sulfate lifetime, SO_2 oxidation rates, and the effect of humidity on the aerosol mass extinction coefficient. Although undoubtedly useful and informative as a basis for advancing a qualitative understanding of processes, these findings make it far from clear that comprehensive modeling of aerosol forcing alone is relevant to the quantification of uncertainty in aerosol forcing.

One advantage of the simple approach adopted here is that, even if one does not accept my arguments, they help identify what would be required for an aerosol forcing to be considerably more negative than about $-1.0\,\mathrm{W\,m^{-2}}$. If, for instance, SO_2 emissions in 1950 relative to 1975 are too large in the estimates by Smith et al. (2011), or if the forcing from aerosol–cloud interactions is for some reason linear in global SO_2, a more negative aerosol forcing becomes plausible. The latter could arise because emissions become increasingly distributed, as two widely separate sources each contributing 50% to the total emissions will, all things else being equal, contribute a greater forcing than a single source producing all of the emissions. The distributed source argument is however most effective in the case of entirely new sources. Emissions of SO_2 by China in 1980, 35 years ago, were still 10% of the global mean. So although emissions there have increased threefold in the past 30 years, it may well be that additional forcing had reached the point of diminishing returns long ago. One way to explore these ideas would be to extend Eq. (1) to incorporate two sources, such that $Q_a = Q_{a,1} + Q_{a,2}$, where rather than interpreting the two sources physically, they are instead used to optimally decompose the spatiotemporal–compositional pattern of aerosol burdens worldwide. Such a model would still be simple enough to be tractable, something that is essential if the ideas are to be held accountable to physical reasoning, but would be able to account for the possibility that the present model [i.e., Eq. (1)] insufficiently considers the effects from changing patterns of emissions.

Irrespective of the ultimate strength of the aerosol forcing, evidence that it has changed over the part of the observational record (the last 30–50 years) most useful for constraining the major terms in Earth's energy budget is scant. This finding alone lends credence to recent, somewhat lower, estimates of the transient climate response (Bengtsson and Schwartz 2013) and suggests that the limitations associated with an insufficiently detailed understanding of aerosol forcing may be less of an obstacle to progress than previously thought.

Acknowledgments. I thank the Max Planck Society for the Advancement of Science for its support for the freedom of scientific research. I also thank the Lorenz Center at MIT for hosting the author during a period of time when some of these ideas were developed, during which time discussions with Kerry Emanuel, Paul O'Gorman, Dan Rothman, and Susan Solomon are gratefully acknowledged. Additional support was provided through funding from the European Union Seventh Framework Programme (FP7/2007-2013) under Grant Agreement 244067. Sandrine Bony, Saskia Brose, Jean-Louis Dufresne, Andrew Gettleman, Stefan Kinne, Nic Lewis, Robert Pincus, Florian Rauser, Hauke Schmidt, Philip Stier, Robert Wood, and several anonymous reviewers are thanked for comments on draft versions of the manuscript. Anders Engström, Seiji Kato, Stefan Kinne, Norman Loeb, and Wenying Su are thanked for supplementary radiative transfer calculations (Kinne) and analysis of the CERES data (Loeb and Kato), and further analysis of the CMIP-CERES models (Engström) to double check the author's work. Jobst

Müße is thanked for sharing his analysis of droplet concentrations from the AEROCOM Phase II models. Luis Kornblueh is thanked for babysitting the hundred historical simulations, and Thomas Schulthess and the Swiss national supercomputing center (CSCS) are thanked for providing access to their facilities for these simulations. I acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and I thank the climate modeling groups (listed in Table A2 of this paper) for producing and making available their model output and the funding agencies and institutions who provided support for coordination and data distribution. The CERES data were obtained from the NASA Langley Research Center Atmospheric Science Data Center. Primary data and scripts used in the analysis and other supplementary information that may be useful in reproducing the author's work are archived by the Max Planck Institute for Meteorology and can be obtained by contacting publications@mpimet.mpg.de.

APPENDIX A

Aerosol–Radiation Interactions (ARI)

The model adopted for $\overline{F}_{\text{ari}}$ is generalized from the framework introduced by Charlson et al. (1992, hereafter C92), which posits that

$$F_{\text{ari}} = -c_r e_r \tau_{550}. \tag{A1}$$

In this expression c_r is a pseudo clear-sky fraction, τ_{550} is the sulfate optical depth in the midvisible wavelength (550 nm) and e_r is a forcing efficiency. The pseudo clear-sky fraction represents what fraction of the sky the aerosol forcing effectively acts over, and thus accounts for the lack of forcing in regions where the aerosol burden is optically masked by the presence of thick clouds, or a bright surface. The actual clear-sky fraction (0.3) is expected to be somewhat smaller than c_r, because for thin clouds the albedo is far from saturated, so that cloud and aerosols have an additive effect. Put differently, for small values of τ the albedo is linear in τ. The forcing efficiency (e_r) can be calculated from radiative transfer theory and describes the change in the top of atmosphere shortwave irradiance from a unit aerosol optical depth. It depends on the atmospheric clear-sky transmissivity, the properties of the aerosol (including how the optical depth varies with wavelength through the visible spectrum), and the background surface albedo. The optical depth measures the extinction of radiation by the aerosol.

The optical depth can be expressed as a product of the mass extinction coefficient k and the aerosol column burden B such that

$$\tau_{550} = kB, \quad \text{where} \quad B = \left(\frac{t_* \frac{3}{2} y}{\Omega}\right) Q_a. \tag{A2}$$

The global burden $B\Omega$, where B is the column burden (in grams per square meter of SO_2) and Ω Earth's surface area, is assumed to be linearly related to SO_2 sources by the yield (oxidation rate) y, which describes the fraction of emitted SO_2 that is converted to sulfate and the sulfate lifetime t_*. The factor of 3/2 accounts for the difference between the molecular weight of SO_2 and the oxidized product, SO_4^{2-} sulfate. It thus assumes that the source is given in units of grams of SO_2. The amount of (radiative) extinction k per unit mass of sulfate can be calculated directly from Mie theory, given an assumed distribution of the aerosol and an ambient relative humidity. It depends on how the aerosol burden is distributed in the vertical, particularly its covariability with humidity upon which k (through the deliquescence effect of the aerosol) is very dependent. Together these expressions imply that

$$F_{\text{ari}} = -e_r k \left(\frac{t_* \frac{3}{2} y}{\Omega}\right) Q_a, \tag{A3}$$

showing that the forcing in a local column is linear in the local source.

To apply this theory globally it is necessary to average over space and time, which because many of the terms can covary both spatially and temporally can introduce spatial and temporal covariances. If a quantity p is the product of a factor s and a variable x, then upon averaging (denoted by overbar, with deviations denoted by a prime) covariances contribute to the mean, that is,

$$\overline{p} = \overline{s}\,\overline{x} + \overline{s'x'}. \tag{A4}$$

To incorporate the effect of covariances I define an effective factor $S = \overline{p}/\overline{x}$, from which it follows that

$$S = \overline{s} + \frac{\overline{s'p'}}{\overline{x}}, \tag{A5}$$

and hence $\overline{p} = S\overline{x}$. This approach is adopted, with the effective yield, lifetime, mass extinction coefficient, radiative efficiency, and pseudo clear-sky fraction denoted by capitalization, when relating the globally averaged forcing to the globally averaged SO_2 source, so that

TABLE A1. Key parameters in determining α in the expression for $\overline{F}_{\rm ari}$. Summarized from results of the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP) and AEROCOM Phase II studies (Schulz et al. 2006; Myhre et al. 2013b).

Parameter	$X_{\rm mean}$	σ_X	$X_{\rm max}$	$X_{\rm min}$	No. of models
Y	0.62	0.15	0.85	0.30	11
T_* (days)	3.8	1.0	5.1	2.3	10
K ($m^2\,g^{-1}$)	11.3	7.0	38.6	4.5	23
$\overline{\tau}_{550}$	0.002	0.0009	0.004	0.0006	30
C_rE_r ($W\,m^{-2}$)	−17.4	5.4	−8.0	−32.0	23
ηC_rE_r ($W\,m^{-2}$)	−12.8	8.0	−3.1	−27.2	25

$$\overline{F}_{\rm ari} = -\left(\frac{C_rE_rKT_*\frac{3}{2}Y}{\Omega}\right)\overline{Q}_a. \qquad \text{(A6)}$$

The five effective parameters that go into the definition of $\overline{F}_{\rm ari}$ are frequently used to interpret the magnitude of $\overline{F}_{\rm ari}$ produced by comprehensive models (e.g., Schulz et al. 2006; Myhre et al. 2013b).

The form for the expression of $\overline{F}_{\rm ari}$ derived above arises from physical considerations and involves relatively few assumptions. Because the effective quantities cannot be measured directly (e.g., in the laboratory) but must be inferred from other measurements spanning a large range of space and time scales, they are uncertain. Values of the parameters derived from comprehensive modeling are provided in Table A1. By comparison, Charlson et al. (1992) estimated an optical depth of 0.04 twice as large as the models, largely due to an overestimate of $\overline{Q}_a$. They further estimated the clear-sky radiative efficiency E_r to be $83\,W\,m^{-2}$ per unit of optical depth as compared to values derived from more detailed radiative transfer calculations of about $25\,W\,m^{-2}$ Their very large (factor of 6) overestimate of the clear-sky forcing was somewhat ameliorated by the assumption of an effective clear-sky fraction of 0.4, as compared to the more conservative 0.6 that arises when aerosol scattering above thin clouds is accounted for.

In deriving the all-aerosol $\overline{F}_{\rm ari}$ as is presented in the manuscript, I assume that the $\overline{F}_{\rm ari}$ from nonsulfate aerosols is proportional, by a factor η, to that produced from sulfate. This assumption has a long history in the literature. It is discussed further in the manuscript, along with supporting evidence below. Given this assumption I arrive at the expression

$$\overline{F}_{\rm ari} = -\alpha\overline{Q}_a, \quad \text{with} \quad \alpha \equiv \eta\left(\frac{C_rE_rKT_*\frac{3}{2}Y}{\Omega}\right), \qquad \text{(A7)}$$

which is used to represent the aerosol–radiation interactions in Eq. (1).

Because estimates of the effective yield and sulfate lifetime are available from relatively few model studies, α in Eq. (1) is estimated from the modeling as $\eta CE\overline{\tau}_{550}/\overline{Q}_a$ for a present-day source of $130\,Tg\,SO_2\,yr^{-1}$. This corresponds to $\alpha = 0.0197\,W\,m^{-2}\,(Tg\,SO_2)^{-1}$. A slightly smaller value, $\alpha = 0.018\,75\,W\,m^{-2}\,(Tg\,SO_2)^{-1}$, is adopted in the fit shown in Fig. 2, and for random parameter draws of the model α^{-1} is assumed to be Gaussian distributed with a mean value and standard deviation of $600 \pm 200\,(W\,m^{-2})^{-1}$ Tg SO_2, corresponding to a central value of $0.0167\,W\,m^{-2}\,(Tg\,SO_2)^{-1}$ and a 2σ range of 0.001 to $0.005\,W\,m^{-2}\,(Tg\,SO_2)^{-1}$.

APPENDIX B

Aerosol–Cloud Interactions (ACI)

Assuming an aerosol perturbation only affects N, then locally the forcing is given by Eq. (3). If all clouds were the same, $I_{\rm cld}$ would adopt values of zero or one depending on whether or not a cloud was present. As stated in the body of the manuscript, cloud macrophysical changes can be accounted for by allowing $I_{\rm cld}$ to adopt other values. This approach projects any variability in E onto $I_{\rm cld}$ and simplifies the expression, at the expense of having to interpret $\overline{I}_{\rm cld}$ as the equivalent "stratocumulus" cloud fraction rather than the actual cloud fraction. To estimate $\overline{F}_{\rm aci}$ I expand the terms in Eq. (3) into a global and annual mean value and a deviation from that value, that is,

$$I \equiv \overline{I}_{\rm cld} + I'_{\rm cld}(x,y,t).$$

If one assumes that third-order terms (the products of three primes) are negligible and that $N'^2 \ll \overline{N}^2$, then it is straightforward to derive the expression for C given by Eq. (5). The covariance terms arise because patterns of forcing $(\delta N)'$ imprint themselves on patterns of clouds, even if $I_{\rm cld}$ is not allowed to directly depend on N. Such correlations were also included in the definition of effective parameters in the expression $\overline{F}_{\rm ari}$, where they play less of a role because the radiative response to a change in an aerosol burden is relatively linear, as long the burden is small and the background is not too bright. For the case of aerosol–cloud interactions these correlation terms are important and invariably act to reduce the net forcing. For instance, larger perturbations in $\delta N'$ are expected over land, where N is already large and in arid conditions, where $I'_{\rm cld}$ is negative.

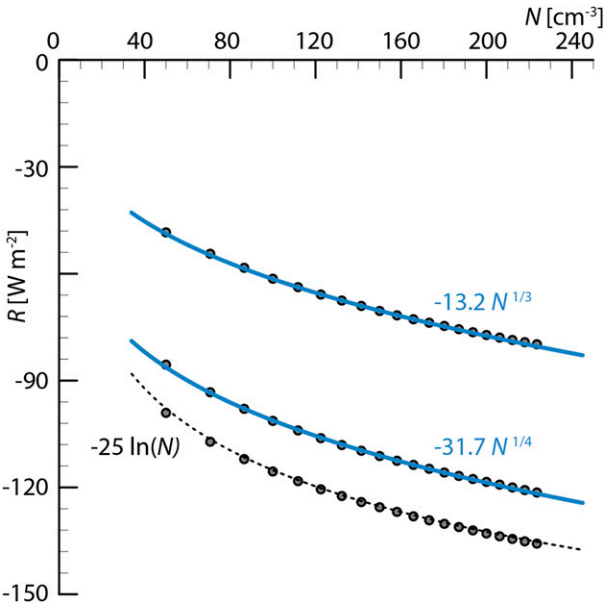

FIG. B1. Cloud radiative effect R for a solar zenith angle of 43.66° assuming a net downward shortwave radiation equal to the tropical (35°S–35°N) average. Shown are radiative transfer calculations (points) for a cloud with liquid water content of 0.35 g kg^{-1} (stronger forcing) and 0.10 g kg^{-1} (weaker forcing) for a cloud inhomogeneity factor of 0.75, as well as a homogeneous cloud with liquid water content of 0.32 g kg^{-1}. Analytic fits to the points assuming a power-law or logarithmic dependence on N are illustrated by the lines.

Charlson et al. (1992) and the subsequent literature often has focused on marine stratocumulus clouds, when considering the possible magnitude of $\overline{F}_{\text{aci}}$. For a typical subtropical stratocumulus layer, analytic arguments can be used to estimate $E = 22\,\text{W m}^{-2}$. Physically the factor E depends on the amount of cloud water, the solar zenith angle, and the net incoming radiation. The value of $22\,\text{W m}^{-2}$ adopted by C92 is consistent with a cloud whose liquid water path is about $65\,\text{g m}^{-2}$, which would correspond to a homogeneous and adiabatic cloud, roughly 300-m deep. Such a cloud is not atypical of stratocumulus forming in a well-mixed boundary layer, such as are often found in eastern boundary current regions of the subtropics (vanZanten et al. 2005; Stevens et al. 2005). In more trade wind–like conditions, or as stratocumulus-topped boundary layers decouple during the day, a thinner cloud and hence smaller value of R would be expected. However, the strength of the forcing from a perturbation in N, which depends on the slope of the curves in Fig. B1, does not change substantially if one assumes a thinner cloud, as the reduction in R is compensated by a greater (power law) dependence on N. Consequently Eq. (3), with $E = 22\,\text{W m}^{-2}$, is a good approximation, even if it implies a too large cloud radiative effect R.

To calculate the effective cloud fraction C radiative transfer calculations were performed (courtesy of S. Kinne) with distributions of present-day clouds using data from the International Satellite Cloud Climatology Project (ISCCP). All liquid clouds were initially assigned an effective radius of 10 μm, which was reduced by 5% to a value of 9.5 μm. So doing results in a globally averaged radiative forcing, $F = -1.52\,\text{W m}^{-2}$ (not including $0.05\,\text{W m}^{-2}$ of offsetting long wave forcing). For a fixed shape of the droplet distribution, the effective radius r_e is proportional to the droplet concentration, such that $r_e \propto N^{-1/3}$ so that the forcing expression can equivalently be written in terms of the effective radius,

$$\overline{F}_{\text{aci}} = \delta R = 3\overline{I}_{\text{cld}} E \frac{\delta r_e}{r_e}. \tag{B1}$$

For uniform perturbations, spatiotemporal covariances are not a factor. From these calculations an effective cloud fraction as $C = 0.46$ is inferred. Repeating the calculations with a smaller 6 μm reference size has only a small effect, slightly reducing the effective cloud fraction. Repeating the calculations and interpreting the results in terms of changes to $\overline{N}$ lead to yet smaller estimates of C, from which the initial upper bound (i.e., one that does not account for covariance terms) of $C = 0.4$, which appears in the main text, is derived.

APPENDIX C

Models, Data, and Methods

For primary data the present study relies on simulations provided by many modeling centers as part of CMIP5 (Taylor et al. 2012). A complete list of the models and the experiments used in this study is provided in Table C1, along with a reference describing each model and its associated experiments (when available).

Radiant energy budgets are taken from the Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) and SYN products (Loeb et al. 2009). Note that Ed2.8 data were mostly used, but compared to Ed2.7 and Ed3.0 data. Different editions of the data did not influence the results. The albedo is constructed from the monthly climatology from 13 years (March 2000 through April 2013) of upward clear sky and downward shortwave irradiances at the top of the atmosphere. The data are processed on the native CERES 1° × 1° latitude–longitude grid, and monthly fluxes are weighted by days per month in forming the long-term average. Land values and monthly values without insolation are masked. The median surface temperature

TABLE C1. CMIP5 models and experiments used in this study.

Model name	AMIP	Hist	Nat	SSTclim	Reference
ACCESS1.0	✓	✓			Bi et al. (2013)
ACCESS1.3		✓			Bi et al. (2013)
BCC_CSM1.1		✓	✓	✓	Xin et al. (2013)
BCC_CSM1.1(m)		✓			Xin et al. (2013)
BNU-ESM	✓	✓			Ji et al. (2014)
CanAM4	✓				von Salzen et al. (2013)
CanESM2		✓	✓	✓	von Salzen et al. (2013)
CCSM4	✓	✓			Meehl et al. (2012)
CESM1(CAM5)	✓	✓			Meehl et al. (2013)
CMCC-CESM		✓			No reference
CMCC-CMS		✓			No reference
CMCC-CM		✓			No reference
CNRM-CM5	✓	✓	✓		Voldoire et al. (2013)
CNRM-CM5.2		✓			Voldoire et al. (2013)
CSIRO Mk3.6.0	✓	✓	✓	✓	Rotstayn et al. (2012)
FGOALS-g2		✓		✓	Li et al. (2013)
GFDL CM3	✓	✓	✓		Donner et al. (2011)
GFDL-ESM2G		✓			Donner et al. (2011)
GFDL-ESM2M		✓			Donner et al. (2011)
GISS-E2-H-CC		✓			Schmidt et al. (2014)
GISS-E2-R	✓	✓			Schmidt et al. (2014)
GISS-E2-R-CC		✓			Schmidt et al. (2014)
HadCM3		✓			Gordon et al. (2000)
HadGEM2-A	✓			✓	Collins et al. (2011)
HadGEM2-AO		✓			Collins et al. (2011)
HadGEM2-CC		✓			Collins et al. (2011)
HadGEM2-ES		✓	✓		Jones et al. (2011)
INM CM4	✓	✓			Volodin et al. (2010)
IPSL-CM5A-LR	✓	✓	✓	✓	Dufresne et al. (2013)
IPSL-CM5A-MR		✓			Dufresne et al. (2013)
IPSL-CM5B-LR	✓	✓			Hourdin et al. (2013)
MIROC5	✓	✓		✓	Watanabe et al. (2010)
MIROC-ESM-CHEM		✓	✓		Watanabe et al. (2011)
MIROC-ESM		✓			Watanabe et al. (2011)
MPI-ESM-LR		✓		✓	Stevens et al. (2013)
MPI-ESM-MR	✓	✓			Giorgetta et al. (2013)
MPI-ESM-P		✓			Giorgetta et al. (2013)
MRI-CGCM3	✓	✓		✓	Yukimoto et al. (2012)
MRI-ESM1		✓			Yukimoto et al. (2012)
NorESM1-M	✓	✓	✓		Bentsen et al. (2013)
NorESM1-ME		✓			Tjiputra et al. (2013)

estimates are from the HadCRUT4.2.0.0 product (Morice et al. 2012). For the aerosol data, use is made of MAC-v1.0 (Kinne et al. 2013), which describes the optical properties of tropospheric aerosols on monthly time scales, discriminated partly by species so as to separate sulfate from other contributions, and with global coverage also on a 1° grid. The climatology is developed from locally sparse, but high-quality, data collected from the AERONET ground-based sun-photometer network, and merged onto complete background maps defined by central data from global aerosol models.

The SO_2 emissions are derived from the tabulated values provided by Smith et al. (2011). The decadal data of anthropogenic source strength taken from their Tables 2 and 3 are plotted in Fig. 1, along with uncertainty estimates. For analytic purposes a curve is fit to the data, and includes all anthropogenic sources, including agriculture and grassland/forest burning, so that in the present paper we set

$$\overline{Q}_a = 2 + 137\left[1 + \exp\left(\frac{1960 - t}{27}\right)\right]^{-1} + 42\exp\left[-\left(\frac{1975 - t}{16}\right)^2\right] + 12\exp\left[-\left(\frac{t - 1915}{25}\right)^2\right] \quad \text{(C1)}$$

with t measuring calendar (Gregorian) year. This fit to the data was developed by eye and is used in the evaluation of $\overline{F}_{aer}$.

In calculating greenhouse gas forcing I used concentrations taken from the data provided by the representative concentration pathways, which were developed for CMIP5 (PRE2005_MIDYR_CONC.DAT). Concentrations for CO_2, CH_4, N_2O, and all gases controlled under the Montreal Protocol (expressed as CFC-12 equivalent concentrations) are converted to a forcing using the simplified expressions provided in Ramaswamy et al. (2001). Stratospheric ozone (a negative forcing) and tropospheric ozone (a somewhat larger positive forcing) are not considered, and assumed to be offset by land-use changes, which are commensurate with the net ozone forcing, but of opposite sign (Myhre et al. 2013a). It is estimated that accounting for these forcings could influence the estimates of the lower bound on $\overline{F}_{aer}$ by 5%–10%, that is, at the level of significance of the estimate (which is given to two significant figures only). To estimate the scaled total aerosol forcing for the Carslaw et al. (2013) study, as shown in Fig. 2, a term had to be added for aerosol–radiation interactions. In doing so I assumed it to be linear in emissions and to contribute 50% of the forcing in the year 2000, as in AR5.

Radiative transfer calculations to estimate parameter values and assumptions in the model for $\overline{F}_{ari}$ (as described below), are performed using the PSRad implementation of RRTM (Pincus and Stevens 2013; Mlawer et al. 1997). Single-column radiation calls were performed using a slightly modified version of the Air Force Geophysics Laboratory tropical sounding. Modifications were made to increase the resolution in the lower troposphere and adjust the thermodynamic profiles consistent with these changes. The temperature profile was modified so that in the layer below 920-hPa temperature decreased following a dry adiabatic lapse rate, and the layer between 920 and 700 hPa had a more moist adiabatic lapse rate of $-6\,K\,km^{-1}$. The humidity in the lower layer increased from 65% at the surface to 85% at 920 hPa. In the upper layer the relative humidity decreased from 70% at cloud base to 50% at 700 hPa. The vertical resolution was specified at 20 hPa in the lower 1 km of the sounding, increasing gradually above that to roughly 1-km-thick levels for pressures lower than 700 hPa. For most of the calculations the aerosol burden was specified as in previous studies (Stier et al. 2013), so that the optical depth was uniformly distributed over the lower 2 km of the model. Sensitivity tests were performed with a more complex vertical structure of the aerosol, based on the climatological distribution from MAC-v1.0. In these calculations the aerosol burden was distributed so that the optical depth in each layer was constant below a height of 1250 m and decreased exponentially above 1250 m with a length scale of 750 m. For some sensitivity studies a background natural aerosol was introduced and characterized by an optical depth of 0.1, a single scattering albedo of 0.97 and an asymmetry factor of 0.7. For all calculations involving sulfate, a wavelength independent asymmetry factor and single scattering albedo of 0.65 and 0.999 999 respectively were specified. The presence of a background aerosol has a minor (5%) effect and slightly reduces the transmissivity. In performing the radiative transfer calculations an effective zenith angle had to be specified. For this the globally and annually averaged insolation-weighted zenith angle of 48.2° (the cosine of which is 0.6667) was used for estimates of global forcing.

In exploring the asymmetry of the albedo in the CERES measurements averages are taken over the region between 25° and 50°N. Over this latitude belt the averaged irradiance is $337.6\,W\,m^{-2}$, and the averaged zenith angle is 49.6°. For calculations related to cloud forcing of tropical clouds, the tropical average for a broad tropical region, equatorward of 35°, was defined. In this region the irradiance weighted zenith angle reduces to 43.66°N and the averaged solar irradiance is $390\,W\,m^{-2}$.

For the simulations with the MPI-ESM, version 1.1 of that model was used, with the 100 ensemble members starting from a preindustrial control simulation. Ensemble members were spawned every 48 years from the control simulation, with the first ensemble member starting from year 48 of the control.

REFERENCES

Bengtsson, L., and S. E. Schwartz, 2013: Determination of a lower bound on Earth's climate sensitivity. *Tellus,* **65B,** 21533, http://dx.doi.org/10.3402/tellusb.v65i0.21533.

Bentsen, M., and Coauthors, 2013: The Norwegian Earth System Model, NorESM1-M—Part 1: Description and basic evaluation of the physical climate. *Geosci. Model Dev.,* **6,** 687–720, doi:10.5194/gmd-6-687-2013.

Bi, D., and Coauthors, 2013: The ACCESS coupled model: Description, control climate and evaluation. *Aust. Meteor. Oceanogr. J.,* **63,** 41–64.

Boucher, O., and T. L. Anderson, 1995: General circulation model assessment of the sensitivity of direct climate forcing by anthropogenic sulfate aerosols to aerosol size and chemistry. *J. Geophys. Res.,* **100** (D12), 26 117–26 134, doi:10.1029/95JD02531.

——, and M. Pham, 2002: History of sulfate aerosol radiative forcings. *Geophys. Res. Lett.,* **29,** 1308, doi:10.1029/2001GL014048.

——, and J. Quaas, 2012: Water vapour affects both rain and aerosol optical depth. *Nat. Geosci.,* **6,** 4–5, doi:10.1038/ngeo1692.

——, and Coauthors, 2013: Clouds and aerosols. *Climate Change 2013: The Physical Science Basis,* T. Stocker et al., Eds., Cambridge University Press, 571–657.

Brasseur, G. P., and E. Roeckner, 2005: Impact of improved air quality on the future evolution of climate. *Geophys. Res. Lett.,* **32,** L23704, doi:10.1029/2005GL023902.

Callendar, G., 1938: The artificial production of carbon dioxide and its influence on temperature. *Quart. J. Roy. Meteor. Soc.,* **64,** 223–240, doi:10.1002/qj.49706427503.

Carslaw, K. S., and Coauthors, 2013: Large contribution of natural aerosols to uncertainty in indirect forcing. *Nature,* **503,** 67–71, doi:10.1038/nature12674.

Charlson, R. J., J. Langner, H. Rodhe, C. B. Leovy, and S. G. Warren, 1991: Perturbation of the northern hemisphere radiative balance by backscattering from anthropogenic sulfate aerosols. *Tellus,* **43A,** 152–163, doi:10.1034/j.1600-0870.1991.00013.x.

——, S. E. Schwartz, J. M. Hales, R. D. Cess, J. A. Coakley, J. E. Hansen, and D. J. Hofmann, 1992: Climate forcing by anthropogenic aerosols. *Science,* **255,** 423–430, doi:10.1126/science.255.5043.423.

Chen, Y. C., M. W. Christensen, L. Xue, A. Sorooshian, G. L. Stephens, R. M. Rasmussen, and J. H. Seinfeld, 2012: Occurrence of lower cloud albedo in ship tracks. *Atmos. Chem. Phys.,* **12,** 8223–8235, doi:10.5194/acp-12-8223-2012.

Christensen, M. W., and G. L. Stephens, 2011: Microphysical and macrophysical responses of marine stratocumulus polluted by underlying ships: Evidence of cloud deepening. *J. Geophys. Res.,* **116,** D03201, doi:10.1029/2010JD014638.

Collins, W. J., and Coauthors, 2011: Development and evaluation of an Earth-system model—HadGEM2. *Geosci. Model Dev.,* **4,** 1051–1075, doi:10.5194/gmd-4-1051-2011.

Déandreis, C., Y. Balkanski, J. L. Dufresne, and A. Cozic, 2012: Radiative forcing estimates of sulfate aerosol in coupled climate–chemistry models with emphasis on the role of the temporal variability. *Atmos. Chem. Phys.,* **12,** 5583–5602, doi:10.5194/acp-12-5583-2012.

Donner, L. J., and Coauthors, 2011: The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3. *J. Climate,* **24,** 3484–3519, doi:10.1175/2011JCLI3955.1.

Dufresne, J. L., and Coauthors, 2013: Climate change projections using the IPSL-CM5 Earth System Model: From CMIP3 to CMIP5. *Climate Dyn.,* **40,** 2123–2165, doi:10.1007/s00382-012-1636-1.

Ekman, A. M. L., 2014: Do sophisticated parameterizations of aerosol–cloud interactions in CMIP5 models improve the representation of recent observed temperature trends? *J. Geophys. Res. Atmos.,* **119,** 817–832, doi:10.1002/2013JD020511.

Fischer, H., D. Wagenbach, and J. Kipfstuhl, 1998: Sulfate and nitrate firn concentrations on the Greenland ice sheet: 2. Temporal anthropogenic deposition changes. *J. Geophys. Res.,* **103** (D17), 21 935–21 942, doi:10.1029/98JD01886.

Flato, G., and Coauthors, 2013: Evaluation of climate models. *Climate Change 2013: The Physical Science Basis,* T. Stocker et al., Eds., Cambridge University Press, 741–866.

Giorgetta, M. A., and Coauthors, 2013: Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *J. Adv. Model. Earth Syst.,* **5,** 572–597, doi:10.1002/jame.20038.

Golaz, J.-C., M. Salzmann, L. J. Donner, L. W. Horowitz, Y. Ming, and M. Zhao, 2011: Sensitivity of the aerosol indirect effect to subgrid variability in the cloud parameterization of the GFDL atmosphere general circulation model AM3. *J. Climate,* **24,** 3145–3160, doi:10.1175/2010JCLI3945.1.

Gordon, C., C. Cooper, C. A. Senior, H. Banks, J. M. Gregory, T. C. Johns, J. F. B. Mitchell, and R. A. Wood, 2000: The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dyn.,* **16,** 147–168, doi:10.1007/s003820050010.

Hansen, J., 2005: Efficacy of climate forcings. *J. Geophys. Res.,* **110,** D18104, doi:10.1029/2005JD005776.

——, M. Sato, and R. Ruedy, 1997: Radiative forcing and climate response. *J. Geophys. Res.,* **102** (D6), 6831–6864, doi:10.1029/96JD03436.

Hoose, C., J. E. Kristjansson, T. Iversen, A. Kirkevåg, O. Seland, and A. Gettelman, 2009: Constraining cloud droplet number concentration in GCMs suppresses the aerosol indirect effect. *Geophys. Res. Lett.,* **36,** L12807, doi:10.1029/2009GL038568.

Hourdin, F., and Coauthors, 2013: LMDZ5B: The atmospheric component of the IPSL climate model with revisited parameterizations for clouds and convection. *Climate Dyn.,* **40,** 2193–2222, doi:10.1007/s00382-012-1343-y.

IPCC, 2013: Summary for policymakers. *Climate Change 2013: The Physical Science Basis,* T. Stocker et al., Eds., Cambridge University Press, 1–29.

Ji, D., and Coauthors, 2014: Description and basic evaluation of BNU-ESM version 1. *Geosci. Model Dev. Discuss.,* **7,** 1601–1647, doi:10.5194/gmdd-7-1601-2014.

Jones, C. D., and Coauthors, 2011: The HadGEM2-ES implementation of CMIP5 centennial simulations. *Geosci. Model Dev.,* **4,** 543–570, doi:10.5194/gmd-4-543-2011.

Kang, S. M., and S.-P. Xie, 2014: Dependence of climate response on meridional structure of external thermal forcing. *J. Climate,* **27,** 5593–5600, doi:10.1175/JCLI-D-13-00622.1.

Kinne, S., and Coauthors, 2013: MAC-v1: A new global aerosol climatology for climate studies. *J. Adv. Model. Earth Syst.,* **5,** 704–740, doi:10.1002/jame.20035.

Klein, S. A., and D. L. Hartmann, 1993: The seasonal cycle of low stratiform clouds. *J. Climate,* **6,** 1587–1606, doi:10.1175/1520-0442(1993)006<1587:TSCOLS>2.0.CO;2.

Kühn, T., A. I. Partanen, A. Laakso, and Z. Lu, 2014: Climate impacts of changing aerosol emissions since 1996. *Geophys. Res. Lett.,* **41,** 4711–4718, doi:10.1002/2014GL060349.

Li, L., and Coauthors, 2013: The flexible global ocean–atmosphere–land system model, grid-point version 2: FGOALS-g2. *Adv. Atmos. Sci.,* **30,** 543–560, doi:10.1007/s00376-012-2140-6.

Loeb, N. G., B. A. Wielicki, D. R. Doelling, G. L. Smith, D. F. Keyes, S. Kato, N. Manalo-Smith, and T. Wong, 2009: Toward optimal closure of the Earth's top-of-atmosphere radiation budget. *J. Climate,* **22,** 748–766, doi:10.1175/2008JCLI2637.1.

McConnell, J. R., and Coauthors, 2007: 20th-century industrial black carbon emissions altered Arctic climate forcing. *Science,* **317,** 1381–1384, doi:10.1126/science.1144856.

Medeiros, B., and B. Stevens, 2011: Revealing differences in GCM representations of low clouds. *Climate Dyn.,* **36,** 385–399, doi:10.1007/s00382-009-0694-5.

——, ——, and S. Bony, 2015: Using aquaplanets to understand the robust responses of comprehensive climate models to forcing. *Climate Dyn.,* **44,** 1957–1977, doi:10.1007/s00382-014-2138-0.

Meehl, G. A., and Coauthors, 2012: Climate system response to external forcings and climate change projections in CCSM4. *J. Climate,* **25,** 3661–3683, doi:10.1175/JCLI-D-11-00240.1.

——, and Coauthors, 2013: Climate change projections in CESM1 (CAM5) compared to CCSM4. *J. Climate,* **26,** 6287–6308, doi:10.1175/JCLI-D-12-00572.1.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-*k* model for the longwave. *J. Geophys. Res.,* **102** (D14), 16 663–16 682, doi:10.1029/97JD00237.

Morice, C. P., J. J. Kennedy, N. A. Rayner, and P. D. Jones, 2012: Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set. *J. Geophys. Res.,* **117,** D08101, 10.1029/2011JD017187.

Murphy, D. M., 2013: Little net clear-sky radiative forcing from recent regional redistribution of aerosols. *Nat. Geosci.,* **6,** 258–262, doi:10.1038/ngeo1740.

——, S. Solomon, R. W. Portmann, K. H. Rosenlof, P. M. Forster, and T. Wong, 2009: An observationally based energy balance for the Earth since 1950. *J. Geophys. Res.,* **114,** D17107, doi:10.1029/2009JD012105.

Myhre, G., and Coauthors, 2013a: Anthropogenic and natural radiative forcing. *Climate Change 2013: The Physical Science Basis,* T. Stocker et al., Eds., Cambridge University Press, 659–740.

——, and Coauthors, 2013b: Radiative forcing of the direct aerosol effect from AeroCom Phase II simulations. *Atmos. Chem. Phys.,* **13,** 1853–1877, doi:10.5194/acp-13-1853-2013.

Nam, C., S. Bony, J. L. Dufresne, and H. Chepfer, 2012: The 'too few, too bright' tropical low-cloud problem in CMIP5 models. *Geophys. Res. Lett.,* **39,** L21801, doi:10.1029/2012GL053421.

Pincus, R., and B. Stevens, 2013: Paths to accuracy for radiation parameterizations in atmospheric models. *J. Adv. Model. Earth Syst.,* **5,** 225–233, doi:10.1002/jame.20027.

Prather, M., G. Flato, P. Friedlingstein, C. Jones, J.-F. Lamarque, H. Liao, and P. Rasch, 2013: Annex II: Climate system scenario tables. *Climate Change 2013: The Physical Science Basis,* T. Stocker et al., Eds., Cambridge University Press, 1395–1445.

Ramaswamy, V., and Coauthors, 2001: Radiative forcing of climate change. *Climate Change 2001: The Scientific Basis,* J. T. Houghton et al., Eds, Cambridge University Press, 349–416.

Rotstayn, L. D., S. J. Jeffrey, M. A. Collier, S. M. Dravitzki, A. C. Hirst, J. I. Syktus, and K. K. Wong, 2012: Aerosol- and greenhouse gas-induced changes in summer rainfall and circulation in the Australasian region: A study using single-forcing climate simulations. *Atmos. Chem. Phys.,* **12,** 6377–6404, doi:10.5194/acp-12-6377-2012.

Schmidt, G. A., and Coauthors, 2014: Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive. *J. Adv. Model. Earth Syst.,* **6,** 141–184, doi:10.1002/2013MS000265.

Schulz, M., C. Textor, and S. Kinne, 2006: Radiative forcing by aerosols as derived from the AeroCom present-day and pre-industrial simulations. *Atmos. Chem. Phys.,* **6,** 5225-5246, doi:10.5194/acp-6-5225-2006.

Sherwood, S. C., S. Bony, O. Boucher, C. Bretherton, P. M. Forster, J. M. Gregory, and B. Stevens, 2015: Adjustments in the forcing–feedback framework for understanding climate change. *Bull. Amer. Meteor. Soc.,* **96,** 217–228, doi:10.1175/BAMS-D-13-00167.1.

Shindell, D. T., 2014: Inhomogeneous forcing and transient climate sensitivity. *Nat. Climate Change,* **4,** 274–277, doi:10.1038/nclimate2136.

——, and Coauthors, 2013: Radiative forcing in the ACCMIP historical and future climate simulations. *Atmos. Chem. Phys.,* **13,** 2939–2974, doi:10.5194/acp-13-2939-2013.

Skeie, R. B., T. K. Berntsen, G. Myhre, K. Tanaka, M. M. Kvalevåg, and C. R. Hoyle, 2011: Anthropogenic radiative forcing time series from pre-industrial times until 2010. *Atmos. Chem. Phys.,* **11,** 11 827–11 857, doi:10.5194/acp-11-11827-2011.

Smith, S. J., J. van Aardenne, Z. Klimont, R. J. Andres, A. Volke, and S. Delgado Arias, 2011: Anthropogenic sulfur dioxide emissions: 1850–2005. *Atmos. Chem. Phys.,* **11,** 1101–1116, doi:10.5194/acp-11-1101-2011.

Sohn, B. J., T. Nakajima, M. Satoh, and H. S. Jang, 2010: Impact of different definitions of clear-sky flux on the determination of longwave cloud radiative forcing: NICAM simulation results. *Atmos. Chem. Phys.,* **10,** 11 641–11 646, doi:10.5194/acp-10-11641-2010.

Stevens, B., and G. Feingold, 2009: Untangling aerosol effects on clouds and precipitation in a buffered system. *Nature,* **461,** 607–613, doi:10.1038/nature08281.

——, and S. E. Schwartz, 2012: Observing and modeling Earth's energy flows. *Surv. Geophys.,* **33,** 779–816, doi:10.1007/s10712-012-9184-0.

——, and Coauthors, 2005: Evaluation of large-eddy simulations via observations of nocturnal marine stratocumulus. *Mon. Wea. Rev.,* **133,** 1443–1462, doi:10.1175/MWR2930.1.

——, and Coauthors, 2013: Atmospheric component of the MPI-M Earth system model: ECHAM6. *J. Adv. Model. Earth Syst.,* **5,** 146–172, doi:10.1002/jame.20015.

Stier, P., and Coauthors, 2013: Host model uncertainties in aerosol radiative forcing estimates: Results from the AeroCom prescribed intercomparison study. *Atmos. Chem. Phys.,* **13,** 3245–3270, doi:10.5194/acp-13-3245-2013.

Storelvmo, T., U. Lohmann, and R. Bennartz, 2009: What governs the spread in shortwave forcings in the transient IPCC AR4 models? *Geophys. Res. Lett.,* **36,** L01806, doi:10.1029/2008GL036069.

Suo, L., O. Helge Otterå, M. Bentsen, Y. Gao, and O. M. Johannessen, 2013: External forcing of the early 20th century Arctic warming. *Tellus,* **65A,** 20578, http://dx.doi.org/10.3402/tellusa.v65i0.20578.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.,* **93,** 485–498, doi:10.1175/BAMS-D-11-00094.1.

Tjiputra, J. F., C. Roelandt, M. Bentsen, D. M. Lawrence, T. Lorentzen, J. Schwinger, Ø. Seland, and C. Heinze, 2013: Evaluation of the carbon cycle components in the Norwegian Earth System Model (NorESM). *Geosci. Model Dev.,* **6,** 301–325, doi:10.5194/gmd-6-301-2013.

vanZanten, M. C., B. Stevens, G. Vali, and D. H. Lenschow, 2005: Observations of drizzle in nocturnal marine stratocumulus. *J. Atmos. Sci.,* **62,** 88–106, doi:10.1175/JAS-3355.1.

Voigt, A., S. Bony, J.-L. Dufresne, and B. Stevens, 2014a: The radiative impact of clouds on the shift of the intertropical convergence zone. *Geophys. Res. Lett.,* **41,** 4308–4315, doi:10.1002/2014GL060354.

——, B. Stevens, J. Bader, and T. Mauritsen, 2014b: Compensation of hemispheric albedo asymmetries by shifts of the ITCZ and tropical clouds. *J. Climate,* **27,** 1029–1045, doi:10.1175/JCLI-D-13-00205.1.

Voldoire, A., and Coauthors, 2013: The CNRM-CM5.1 global climate model: Description and basic evaluation. *Climate Dyn.,* **40,** 2091–2121, doi:10.1007/s00382-011-1259-y.

Volodin, E. M., N. A. Dianskii, and A. V. Gusev, 2010: Simulating present-day climate with the INMCM4.0 coupled model of the atmospheric and oceanic general circulations. *Izv. Atmos. Ocean. Phys.,* **46,** 414–431, doi:10.1134/S000143381004002X.

von Salzen, K., and Coauthors, 2013: The Canadian Fourth Generation Atmospheric Global Climate Model (CanAM4).

Part I: Representation of physical processes. *Atmos.–Ocean,* **51,** 104–125, doi:10.1080/07055900.2012.755610.

Watanabe, M., and Coauthors, 2010: Improved climate simulation by MIROC5: Mean states, variability, and climate sensitivity. *J. Climate,* **23,** 6312–6335, doi:10.1175/2010JCLI3679.1.

Watanabe, S., and Coauthors, 2011: MIROC-ESM 2010: Model description and basic results of CMIP5-20c3m experiments. *Geosci. Model Dev.,* **4,** 845–872, doi:10.5194/gmd-4-845-2011.

Wilcox, L. J., E. J. Highwood, and N. J. Dunstone, 2013: The influence of anthropogenic aerosol on multi-decadal variations of historical global climate. *Environ. Res. Lett.,* **8,** 024033, doi:10.1088/1748-9326/8/2/024033.

Wood, R., 2012: Stratocumulus clouds. *Mon. Wea. Rev.,* **140,** 2373–2423, doi:10.1175/MWR-D-11-00121.1.

Wyant, M. C., and Coauthors, 2015: Global and regional modeling of clouds and aerosols in the marine boundary layer during VOCALS: The VOCA intercomparison. *Atmos. Chem. Phys.,* **15,** 153–172, doi:10.5194/acp-15-153-2015.

Xin, X., T. Wu, J. Li, Z. Wang, W. Li, and F. Wu, 2013: How well does BCC CSM1.1 reproduce the 20th century climate change over China? *Atmos. Oceanic Sci. Lett.,* **6,** 21–26.

Yukimoto, S., and Coauthors, 2012: A new global climate model of the Meteorological Research Institute: MRI-CGCM3—Model description and basic performance. *J. Meteor. Soc. Japan,* **90A,** 23–64, doi:10.2151/jmsj.2012-A02.

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