

Reply to “Comment on ‘Rethinking the Lower Bound on Aerosol Radiative Forcing’”

BJORN STEVENS AND STEPHANIE FIEDLER

Max Planck Institute for Meteorology, Hamburg, Germany

(Manuscript received 27 January 2017, in final form 28 May 2017)

ABSTRACT

Kretzschmar et al., in a comment in 2017, use the spread in the output of aerosol–climate models to argue that the models refute the hypothesis (presented in a paper by Stevens in 2015) that for the mid-twentieth-century warming to be consistent with observations, then the present-day aerosol forcing, F_{aer} , must be less negative than -1 W m^{-2} . The main point of contention is the nature of the relationship between global SO_2 emissions and F_{aer} . In contrast to the concave (log-linear) relationship used by Stevens and in earlier studies, whereby F_{aer} becomes progressively less sensitive to SO_2 emissions, some models suggest a convex relationship, which would imply a less negative lower bound. The model that best exemplifies this difference, and that is most clearly in conflict with the hypothesis of Stevens, does so because of an implausible aerosol response to the initial rise in anthropogenic aerosol precursor emissions in East and South Asia—already in 1975 this model’s clear-sky reflectance from anthropogenic aerosol over the North Pacific exceeds present-day estimates of the clear-sky reflectance by the total aerosol. The authors perform experiments using a new (observationally constrained) climatology of anthropogenic aerosols to further show that the effects of changing patterns of aerosol and aerosol precursor emissions during the late twentieth century have, for the same global emissions, relatively little effect on F_{aer} . These findings suggest that the behavior Kretzschmar et al. identify as being in conflict with the lower bound in Stevens arises from an implausible relationship between SO_2 emissions and F_{aer} , and thus provides little basis for revising this lower bound.

Stevens (2015, hereinafter S15) used three lines of reasoning to argue that present-day effective aerosol radiative forcing (F_{aer}) is very likely less negative than -1 W m^{-2} . The most quantitative bound arose from the logic that if one wishes to maintain that some component of the warming in the first half of the twentieth century was anthropogenic in origin, then this bounds F_{aer} to have a smaller magnitude than other positive forcings over the same time period. By adopting a simple model that relates F_{aer} to aerosol-precursor emissions, it becomes possible to use this constraint to bound present-day forcing, given knowledge of the precursor and present-day emissions. The

application of this approach to global forcing and global temperature yielded a lower bound on present-day F_{aer} of -1.3 W m^{-2} , a finding corroborated by subsequent studies using different approaches (Rotstayn et al. 2015). Assuming that the arguments also hold on the hemispheric scale yielded the tighter lower bound of -1 W m^{-2} . This less negative bound on aerosol radiative forcing was shown by S15 to be consistent both with the present understanding applied to very simple models (of the kind once used to argue for much more negative forcings) and with the absence of a strong aerosol signal in observations of the hemispheric contrast of clear-sky surface albedo over the ocean.

Motivated by the arguments of S15, Kretzschmar et al. (2017, hereafter K17) analyze simulations from climate models participating in phase 5 of the Coupled Model Intercomparison Project (CMIP5). By regressing global and hemispheric temperature trends against present-day F_{aer} they show that, consistent with the arguments of S15, most models (36 of 42 simulations analyzed by K17)

Denotes content that is immediately available upon publication as open access.

Corresponding author: Bjorn Stevens, bjorn.stevens@mpimet.mpg.de

DOI: 10.1175/JCLI-D-17-0034.1

© 2017 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

underestimate the observed warming in the 30-yr period between 1920 and 1950, and that models with the most negative F_{aer} most underestimate the observed Northern Hemispheric warming between 1860 and 1950. Although generally producing less warming than has been observed, 30-yr temperature trends between 1920 and 1950 show a weaker relationship to F_{aer} , as is to be expected given the larger internal variability on shorter time scales. Nonetheless, K17 interpret their results as being inconsistent with the arguments of S15, opening the door to the idea of a more negative F_{aer} . Their interpretation is based on two findings: 1) the tendency of CMIP5 models with F_{aer} less negative than -0.8 W m^{-2} to warm more in the Northern Hemisphere than is observed during the 1860–1950 period, and 2) the tendency for some models with F_{aer} more negative than -1 W m^{-2} to capture the observed first century of Northern Hemispheric warming. The latter, their analysis suggests, is a result of aerosol forcing being less efficient (as measured by the ratio of F_{aer} to the global emissions) in the first half of the century than in the second half. Disproportionate warming in a period where the atmosphere is more aerosol laden has long been precluded by the argument that a more pristine atmosphere is more susceptible to anthropogenic forcing from aerosol–cloud interactions (Boucher and Pham 2002; Carslaw et al. 2013). This staple of aerosol research is formally incorporated in S15’s model, which constructs F_{aer} from two terms: one that is linear in emissions, consistent with understanding of aerosol–radiation interactions (whose forcing we denote by F_{ari}), and the other logarithmic in emissions, consistent with understanding of aerosol–cloud interactions, whose forcing is denoted as F_{aci} .

The idea (arising from K17’s first finding) that greater than observed Northern Hemispheric warming arising with an F_{aer} less negative than the S15 lower bound somehow implies that the forcing must be more negative is incorrect. As long as the net forcing over some time interval is positive, it is not possible to separate the response to forcing (i.e., the model sensitivity) from the forcing over that interval. Indeed S15’s central argument was never that F_{aer} correlates with the temperature response; rather, that *the presumption that some portion of an observed temperature trend is forced implies that the forcing shares the same sign as the trend, thereby constraining the forcing.*

S15’s arguments thus really only apply to models with F_{aer} robustly less than -1 W m^{-2} , of which (based on K17’s analysis) there are three: GFDL-CM3, HadGEM2-A, and CSIRO-Mk3.6.0. Two of these models, GFDL-CM3 and HadGEM2-A, have ensemble mean 1860–1950 warming in the Northern Hemisphere less than observed, consistent with S15. The third model, CSIRO-Mk3.6.0,

TABLE 1. Differences between clear-sky reflected shortwave irradiance over the ocean between 55°N and 55°S.

Model	R_{1950}	R_{1975}	R_{2000}
CSIRO-Mk3.6.0	0.24	1.34	1.20
HadGEM2-ESM	0.27	0.86	0.92
GFDL-CM3	0.29	1.20	1.18

appears to robustly contradict the reasoning of S15, this being K17’s second finding.

The disagreement, and this is really the heart of the matter, arises because the CSIRO-Mk3.6.0 model (and to a lesser degree some other models analyzed by K17) predict a convex¹ relationship between estimates of forcing and SO_2 emissions, whereas the model of S15 follows earlier studies (Boucher and Pham 2002; Carslaw et al. 2013) in assuming that this relationship is concave. If the convex relationship is correct, then this would mean that the scaling of a midcentury F_{aer} constraint to the present day would imply a less negative (closer to -1.5 W m^{-2}) lower bound than was posited by S15. Notwithstanding that this in itself would still imply a considerable reduction of the uncertainty in aerosol forcing, the question arises as to whether there are good reasons to take the model output analyzed by K17 at face value, and thereby reject S15’s less negative lower bound.

The idea implicit in K17’s interpretation of the model output is that late-century (Asian dominated) emissions project more strongly onto F_{aer} than do early-century (predominantly European and North American) emissions, despite the latter having occurred upon a more pristine background. This could happen for one of three reasons, which we formulate as hypotheses: Asian emissions of SO_2 are exceptionally effective at covering the North Pacific with aerosol (H1), clouds over the North Pacific are exceptionally susceptible to aerosol perturbations (H2), or early twentieth-century emissions were substantially more absorbing (H3). We evaluate H1 and H2 below and show that while they explain why the models behave as they do, this aspect of the model behavior is not consistent with observations on the one hand, or physical understanding on the other. Hence we see no compelling reason to revise the S15 lower bound on F_{aer} .

To test H1 we estimate the anthropogenic aerosol burden at different time periods, following methods developed in S15. That is, we calculate the difference in the reflected clear-sky shortwave irradiance, over the

¹ By convex we mean upward concavity, so that when $-F_{\text{aer}}$ is plotted against SO_2 emissions the slope is increasing, i.e., greater at high (late twentieth century) emission levels than it is at low (early twentieth century) emission levels.

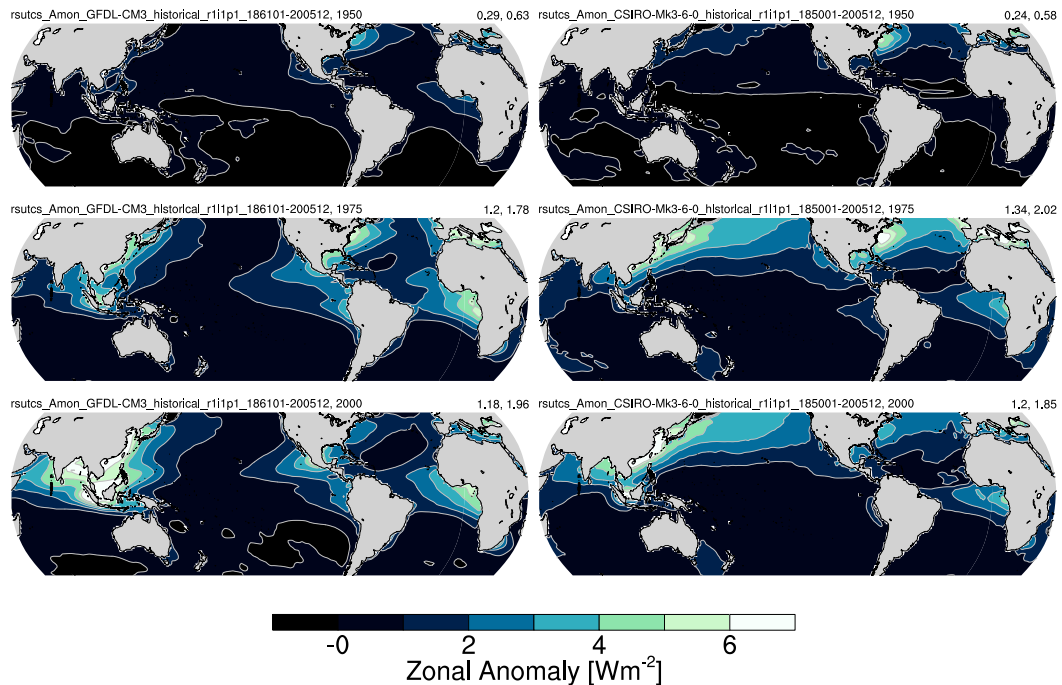


FIG. 1. Patterns of R for three different time periods—from top to bottom, R_{1950} , R_{1975} , and R_{2000} —and two different models: (left) GFDL-CM3 and (right) CSIRO-Mk3.6.0. These two models were chosen because of their very different evolution of F_{aer} in K17. For each year indicated the pattern is constructed for the 10-yr period centered on that year, so that R_{1950} is constructed from reflected shortwave irradiances over the period between 1945 and 1955. The average of R equatorward of 55° for the globe and the Northern Hemisphere alone is indicated on the upper right of each figure panel, in units of W m^{-2} .

ocean, between a given period and a preindustrial period, here taken to be a period between 1861 and 1869 with relatively little volcanic activity. To minimize possible effects from residual volcanic aerosol, and similar to S15, we next calculate the anomaly in this quantity relative to the lowest vigintile (5%) along a line of absolute latitude. To avoid statistical outliers arising from changes in sea ice we limit our analysis to latitudes equatorward of 50° . The resultant quantity, which we denote by R , provides a simple measure of a model's anthropogenic aerosol forcing in a given year (denoted by subscript). Indeed, despite good reason to be skeptical of the F_{aer} estimates,² R_{1950}/R_{2000} varies between 0.2 and 0.3 in the three models analyzed (Table 1 herein), consistent with K17 (their Table 1).

The most striking finding of this analysis is the strong increase in R_{1975} as compared to R_{1950} . This is particularly true of the CSIRO-Mk3.6.0 model output (Fig. 1) but is also evident in other models (e.g., K17) (Fig. 2). Between 1950 and 1975, clear-sky reflected shortwave irradiance anomalies from anthropogenic aerosols

exceed 2 W m^{-2} over most of the North Pacific, but change little thereafter. By contrast the change in R for the GFDL model is much more confined to the continents, and continues to increase with increasing South and East Asian emissions after 1975. Over the Northern Hemisphere extratropics, R_{1975} calculated from the CSIRO-Mk3.6.0 model output is so large that it exceeds the present-day signal from the total aerosol, measured by CERES. Given that globally the total anthropogenic aerosol is estimated to be about 20% of the total aerosol burden (Myhre et al. 2013), and that Asian deserts are a significant source of natural aerosol, it seems that either the CERES data must be in error or the CSIRO-Mk3.6.0 model greatly overestimates the amount of remote aerosol attributable to the initial increase in South and East Asian anthropogenic activity. CALIPSO retrievals (not shown), which are even less susceptible to cloud biases than is our analysis of the CERES data, show a similarly muted signal of aerosol optical depth over the North Pacific and give us confidence in the CERES data. This leads us to conclude that it is the CSIRO-Mk3.6.0 model output that is physically implausible.

Even if the CSIRO-Mk3.6.0 model is not fit for the purpose at hand, it need not call into question H2, namely that clouds over the North Pacific are exceptionally

² K17's F_{aer} estimates include changes from surface albedo, and are made uncertain by interannual variability in top-of-atmosphere irradiances (Fiedler et al. 2017).

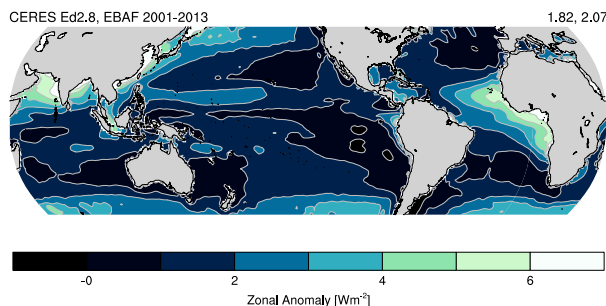


FIG. 2. Zonal anomaly in reflected clear-sky solar irradiance over the ocean from CERES. The anomaly is computed from the lowest vigintile at the absolute latitude. This quantity differs from R in that it includes the natural aerosol, which at high latitudes over the Pacific and at low latitudes over the Atlantic is dominated by mineral dust. The average of R equatorward of 55° for the globe and the Northern Hemisphere alone is indicated on the upper right of each figure panel, in units of W m^{-2} . Original data were smoothed before contouring with two applications of a nine-point smoothing.

susceptible to aerosol perturbations. To test this hypothesis we use the multiple plume representation of aerosol forcing, MACv2-SP, developed by Stevens et al. (2017). This approach relaxes some of the assumptions of S15 but still allows for explicit control over the pattern of aerosol forcing and the relative contributions of F_{ari} versus F_{aci} to the net forcing. Fiedler et al. (2017) compare F_{aer} in 1975, with its value in 2005. In 1975 emissions of aerosols and aerosol precursors in the Atlantic sector are estimated to be more than a factor of 3 larger than those from South and East Asia. By 2005 the situation had reversed, with emissions from South and East Asia becoming threefold larger than emissions from North America and Europe (Stevens et al. 2017). In experiments with ECHAM6.3 wherein F_{aer} was very carefully controlled, this factor of 10 swing in the relative importance of North American/European versus Asian emissions had, however, no discernible effect on the emission weighted value of F_{aer} (Fig. 3). This finding is consistent with the assumptions adopted by S15. It is also consistent with more comprehensive and controlled studies, as Carslaw et al. (2013) also show no change in F_{aci} over a 30-yr period between 1980 and 2000 (see their extended data in Table 4 therein.).

Notwithstanding the inference that shifting patterns of aerosol loading have a small influence on the net forcing, it could be argued that a more discernible effect of the pattern of emissions could arise if aerosol–cloud interactions were a substantially larger fraction of the total forcing than implied by the Stevens et al. (2017) climatology. To test this idea we artificially enhanced the potency of aerosol–cloud interactions in MACv2-SP. Instead of contributing commensurately with aerosol–radiation interactions to F_{aer} , aerosol–cloud interactions were modified to have a threefold larger contribution,

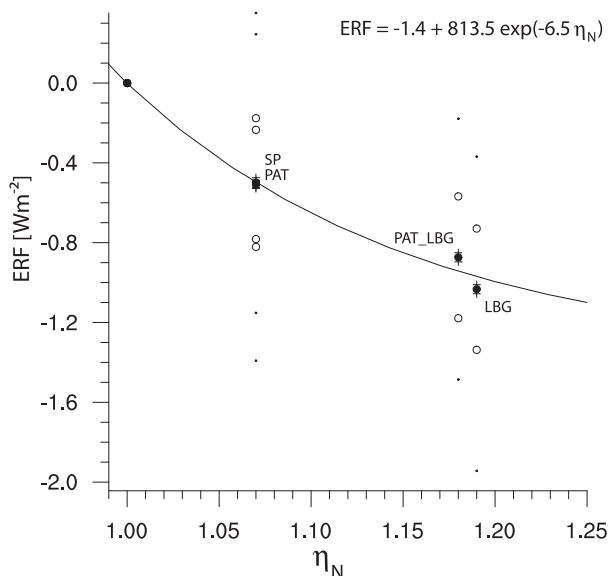


FIG. 3. Effective aerosol radiative forcing, F_{aer} , from ECHAM6.3 plotted by the globally averaged cloud-droplet population density enhancement factor based on the simple-plume aerosol climatology, MACv2-SP. F_{aer} is estimated from an ensemble of three 10-yr simulations for an indicated period and compared to an ensemble of six control runs. Thus for each of the four experiments 180 years of forcing estimates are available. Shown are the mean (filled circle), standard error (horizontal lines), standard deviation of yearly estimates (open circles), and the extremes of annual mean estimates in ERF (points). The four experiments compare the basic climatology, denoted SP, with the same forcing for 1975 (PAT) for which global emissions were about the same as at present. In addition sensitivity experiments are performed in which the background effect is greatly reduced (LBG), yielding a much stronger cloud radiative effect (LBG, and PAT-LBG); for these latter experiments the pattern of the forcing is discernible, but not large (see also Fiedler et al. 2017).

resulting in a $F_{\text{aer}} \approx -1 \text{ W m}^{-2}$ for the present-day anthropogenic aerosol distribution. Maintaining roughly the same global aerosol burden, but instead distributing it following the 1975 pattern of aerosol loading, yields a less negative F_{aer} : -0.87 W m^{-2} . This modest (15%) difference can be expected to be sensitive to the representation of clouds in ECHAM6.3, so that it could be larger. This suggests that models with what we believe to be unrealistically strong aerosol–cloud interactions could also give too much weight to late-twentieth-century emissions, thereby causing the convexity in the relationship between SO_2 emissions and F_{aer} . This is by no means an idiosyncratic interpretation. A recent community assessment addressing this question has revised the magnitude of the aerosol forcing since 1990 to a substantially less negative value, in effect also arguing that the degree of concavity seen in the models analyzed by K17 is implausible (Myhre et al. 2017).

We are not adverse to the idea that F_{aer} may be more negative than the lower bound of S15, possibly for reasons already stated in that paper. We are averse to the idea that climate models, which have gross and well-documented deficiencies in their representation of aerosol–cloud interactions (cf. Boucher et al. 2013), provide a meaningful quantification of forcing uncertainty. Surely after decades of satellite measurements, countless field experiments, and numerous finescale modeling studies that have repeatedly highlighted basic deficiencies in the ability of comprehensive climate models to represent processes contributing to atmospheric aerosol forcing, it is time to give up on the fantasy that somehow their output can be accepted at face value. If progress is to be made in narrowing the bounds on aerosol forcing, it will arise not through a reduction of model spread—which fails to penalize the most erroneous models—but rather by developing and testing specific physical hypotheses underlying a statement for F_{aer} exceeding a given bound.

Acknowledgments. We thank Stefan Kinne for valuable and critical discussions of the K17 comment and our reply. We also would like to acknowledge the collegiality of the authors of K17, and thank them for taking the time to write their comment; formal scientific exchanges are increasingly rare, a trend which belies their value. The authors acknowledge the generous and unfettered support of the Max Planck Society. Use of the supercomputer facilities at the Deutsches Klimarechenzentrum (DKRZ) is acknowledged as is funding from the FP7 project BACCHUS (No. 603445).

REFERENCES

- Boucher, O., and M. Pham, 2002: History of sulfate aerosol radiative forcings. *Geophys. Res. Lett.*, **29**, 1308, doi:10.1029/2001GL014048.
- , and Coauthors, 2013: Clouds and aerosols. *Climate Change 2013: The Physical Science Basis*. T. F. Stocker et al., Eds., Cambridge University Press, 571–657.
- Carlaw, K. S., and Coauthors, 2013: Large contribution of natural aerosols to uncertainty in indirect forcing. *Nature*, **503**, 67–71, doi:10.1038/nature12674.
- Fiedler, S., B. Stevens, and T. Mauritsen, 2017: On the sensitivity of anthropogenic aerosol forcing to model-internal variability and parameterizing a Twomey effect. *J. Adv. Model. Earth Syst.*, doi:10.1002/2017MS000932, in press.
- Kretzschmar, J., M. Salzmann, J. Mülmenstädt, O. Boucher, and J. Quaas, 2017: Comment on “Rethinking the lower bound on aerosol radiative forcing.” *J. Climate*, **30**, 6579–6584, doi:10.1175/JCLI-D-16-0668.1.
- Myhre, G., and Coauthors, 2013: 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.
- , and Coauthors, 2017: Multi-model simulations of aerosol and ozone radiative forcing due to anthropogenic emission changes during the period 1990–2015. *Atmos. Chem. Phys.*, **17**, 2709–2720, doi:10.5194/acp-17-2709-2017.
- Rotstayn, L. D., M. A. Collier, D. T. Shindell, and O. Boucher, 2015: Why does aerosol forcing control historical global-mean surface temperature change in CMIP5 models? *J. Climate*, **28**, 6608–6625, doi:10.1175/JCLI-D-14-00712.1.
- Stevens, B., 2015: Rethinking the lower bound on aerosol radiative forcing. *J. Climate*, **28**, 4794–4819, doi:10.1175/JCLI-D-14-00656.1.
- , S. Fiedler, S. Kinne, K. Peters, S. Rast, J. Müsse, S. J. Smith, and T. Mauritsen, 2017: MACv2-SP: A parameterization of anthropogenic aerosol optical properties and an associated Twomey effect for use in CMIP6. *Geosci. Model Dev.*, **10**, 433–452, doi:10.5194/gmd-10-433-2017.