CORRESPONDENCE

Comments on "Rethinking the Lower Bound on Aerosol Radiative Forcing"

BEN B. B. BOOTH, GLEN R. HARRIS, AND ANDY JONES
Met Office Hadley Centre, Exeter, United Kingdom

LAURA WILCOX
National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Reading, United Kingdom

MATT HAWCROFT
College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, United Kingdom

KEN S. CARSLAW
School of Earth and Environment, University of Leeds, Leeds, United Kingdom

(Manuscript received 7 June 2017, in final form 14 August 2018)

1. Introduction

Stevens (2015, hereinafter S15) used energy balance arguments to estimate a lower limit on real-world aerosol forcings. The essence of this argument is that we expect any externally forced component of the warming between preindustrial and 1950 to have been positive. Therefore we would expect the sign of the corresponding net external forcing to also be positive. S15 uses simple global forcing–emission relationships and historical emission changes to show that large-magnitude present-day aerosol forcing would not be consistent with a 1950 positive net forcing. This analysis predicts that negative present-day aerosol forcings exceeding -1.3 or $-1.0\,\mathrm{W\,m^{-2}}$ can be ruled out based on either 1950 global or Northern Hemispheric (NH) net energy balance, respectively. However, this argument is inconsistent with the warming in available CMIP5 simulations, which brings into question whether such an analysis does indeed imply a constraint on the real world. Out of the 10 CMIP5 simulations for which present-day aerosol forcing estimates are available, six simulate aerosol forcing equal to or larger in magnitude than $-1.0\,\mathrm{W\,m^{-2}}$ and three simulate it equal to or greater than $-1.3\,\mathrm{W\,m^{-2}}$, yet all reproduce a global warming trend, and almost all predict a positive NH trend (see Table 1). Understanding why S15's energy balance analysis is not a good guide of the CMIP5 response is not straightforward. However, we have identified several factors in the S15 analysis that would provide partial explanations. These are 1) the degree of linearity of global aerosol forcing and 2) limitations of the regional energy budget analysis. We also identify two other aspects of the analysis where plausible alternative choices would lead to different constraints on the lower limit of real-world aerosol forcing: 3) past aerosol emissions and 4) choice of analysis period. The impact of adopting these alternative assumptions, in the S15 methodology, suggests that any real-world aerosol forcing constraint is likely to be considerably weaker than the S15 headline results.

We have used a similar simple global forcing model [which is a component of the simple climate model documented in Harris et al. (2013)] to that employed in S15, with which we have been able to replicate the S15 global analysis. There are some differences between the two model setups: for example, we account for ozone, volcanic, and solar forcings whereas S15 does not, and we use an 1860 baseline compared to S15's late 1700 baseline.

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

Corresponding author: Ben B. B. Booth, ben.booth@metoffice.gov.uk

DOI: 10.1175/JCLI-D-17-0369.1

TABLE 1. The global and NH preindustrial to 2000 aerosol forcing (diagnosed from those models reporting sstClim and sstClimAerosol CMIP5 experiments) and global and NH temperature changes between the mean temperature in the 1860 and 1950 decades from the ensemble mean response of the CMIP5 historical simulations [in contrast with Kretzschmar et al. (2017), who regressed off trends through the time series]. An equivalent estimate from the observed mean estimate (HadCRUT4; Morice et al. 2012) is also given. Stevens (2015) implies that present-day aerosol forcing larger than $-1.3\,\mathrm{W\,m^{-2}}$ (boldface) or $-1.0\,\mathrm{W\,m^{-2}}$ (boldface and italics) should be unable to reproduce warming up to 1950, globally or in the NH, respectively.

CMIP5 model	Global forcing ($\mathrm{W\,m^{-2}}$)	NH forcing ($\mathrm{W\,m^{-2}}$)	Global 1860s to 1950s temperature change (K)	NH 1860s to 1950s temperature change (K)
GFDL CM3	**−1.6**	−2.4	**0.16**	***0.00***
CSIRO Mk3.6.0	**−1.4**	−1.7	**0.33**	***0.26***
MIROC5	**−1.3**	−2.1	**0.30**	***0.29***
HadGEM2-A/HadGEM2-ES	***−1.2***	−2.0	0.16	***0.14***
MRI-CGCM3	***−1.1***	−1.4	0.16	***0.26***
NorESM1-M	***−1.0***	−1.4	0.12	***0.11***
CanESM2	−0.9	−1.3	0.26	0.19
MPI-ESM-LR	−0.4	−0.3	0.23	0.31
FGOALS-s2	−0.4	−0.7	0.53	0.56
BCC-CSM1.1	−0.4	−0.7	0.34	0.36
Observations (HadCRUT4)			**0.28**	**0.30**

The two model representations otherwise agree on the general structural form. We find small differences in the global constraints when adopting the same assumptions as in S15 (our global lower limit of aerosol forcing is $-1.4\,\mathrm{W\,m^{-2}}$, compared to $-1.3\,\mathrm{W\,m^{-2}}$ in S15), implying that the impact of differences in the simple models is likely to be minor. What this replication enables us to do is assess the robustness of S15's analysis to a number of assumptions in the method.

Sections 2–5 discuss the four factors identified in the first paragraph of this section, while section 6 provides our outlook on the potential for requiring net positive 1950 energy balance to constrain the range of real-world aerosol forcing.

2. Is global aerosol forcing linear with emissions?

Aerosol–cloud forcing (indirect) effects are locally nonlinear, with stronger radiative responses to aerosol concentrations in cleaner conditions, but this response weakens as the as the background becomes increasingly polluted (Twomey and Squires 1959). Key to the S15 analysis is the amount of aerosol forcing realized by 1950 and the assumption that the same nonlinearity applies globally. However, we show here, in an earlier version of HadGEM2-A (Fig. 1), that the global mean forcing can be remarkably linear with emissions. These new estimates are important because global aerosol forcing through the twentieth century has not previously been published. The suggestion from Fig. 1 is that models explicitly representing indirect aerosol effects can reproduce global forcing that is considerably more linear with aerosol changes than would be expected from the documented nonlinear response in regional changes. More work is needed to understand the factors that lead to this linearity, but they are likely to be linked to the global averaging of regions where the radiative response is increasingly buffered and more sensitive clean/pristine regions that are progressively affected by new emissions. If we accept that the global radiative forcing could respond linearly to aerosol emissions, then this affects any limit implied from a simple energy balance constraint. Repeating the simple model framework outlined in S15 with a linear global aerosol forcing leads to a larger negative aerosol forcing that can be considered consistent with positive net 1950 forcing (2005 values up to $-1.6\,\mathrm{W\,m^{-2}}$; Fig. 1b). This revised energy balance constraint (which rules out aerosol forcings of $-1.6\,\mathrm{W\,m^{-2}}$ and larger) is interesting, but it still does not explain what we see in the CMIP5 ensemble. GFDL-CM3 simulates a net $-1.6\,\mathrm{W\,m^{-2}}$ aerosol forcing but simulates a forced global temperature rise where the energy balance constraint suggests that there should be negligible warming at best.

As an aside it is worth noting that global linearity of forcing to emissions does not imply that the forcing is insensitive to the preindustrial aerosol state. This is because the global linearity emerges from the aggregate of many regional scale aerosol emission plumes, where the background state determines the nonlinear forcing response to aerosol emission (Twomey and Squires 1959; Carslaw et al. 2013). In aggregating many regional plumes the nonlinearity is averaged out but the sensitivity to the background states remains.

3. Hemispheric constraint?

S15 use a similar argument that the sign of the net forcing must match the sign of the forced temperature

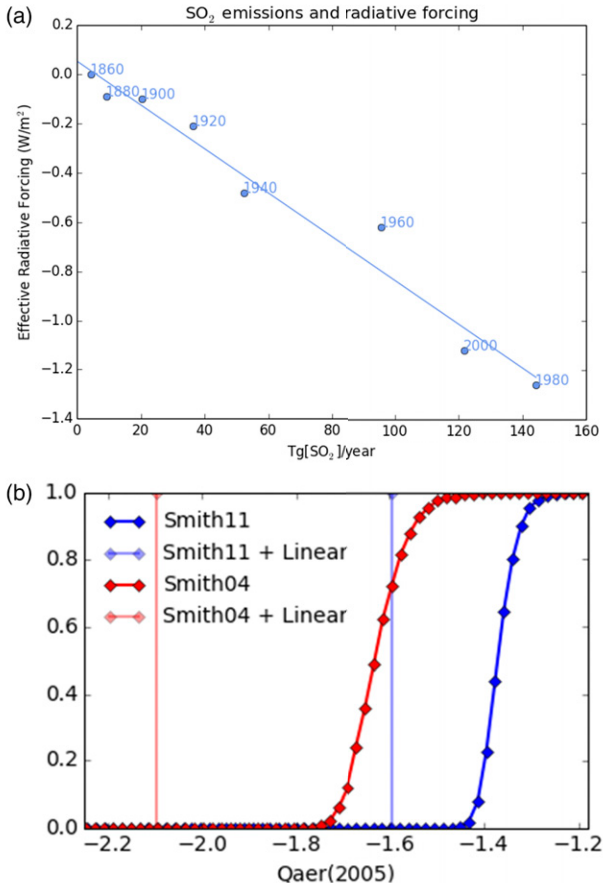

FIG. 1. (a) Effective radiative forcing relative to SO_2 emission changes, calculated from time slices for a range of historical periods for an early development version of HadGEM2A that uses both a model configuration and an SO_2 aerosol emission dataset that predates CMIP5. (b) Probability that the net forcing over the globe between 1850 and 1950 is positive as a function of 2005 aerosol forcing (*x* axis), where other simple model parameters are randomly sampled (as per S15). The blue cumulative distribution function (CDF) uses S15 prior parameter ranges and sulfur emissions from Smith et al. (2011). The red CDF reproduces this, but replaces the sulfur emissions with those from Smith et al. (2004). The vertical blue and red lines represent the constraint from these two emission datasets, respectively, if the aerosol forcing response to emissions can be considered linear.

trend in the NH, in isolation, to come to a stronger constraint that rules out present-day forcings more negative than $-1.0\,\mathrm{W\,m^{-2}}$. However, as Kretzschmar et al. (2017) identified, this constraint does not match NH forcing–temperature relationships seen in CMIP5. Of the six simulations with stronger aerosol forcing than S15's $-1.0\,\mathrm{W\,m^{-2}}$ limit, none produce the forced NH cooling signal up to 1950 (Table 1) implied by S15's analysis. While it may not be the whole story, we identify in our comment a key problem with S15's NH conceptual framework. The S15 constraint is based on the underlying assumption that the sign of hemispheric forcing must match the sign of any forced temperature trend. The problem is that if the other hemisphere were experiencing stronger positive forcing (due to weaker aerosol forcing during the same period) then the resulting cross-equatorial transport of energy to the NH would also tend to warm this hemisphere. Estimates of cross-equatorial energy transfer in response to asymmetric forcings suggest that this heat transfer would be expected to represent a substantial fraction of the Northern and Southern Hemispheric forcing difference, as is evident in both observations (Loeb et al. 2016) and physically based models (Kay et al. 2016; Haywood et al. 2016; Mechoso et al. 2016; Hawcroft et al. 2017). The sign of forced temperature change cannot therefore be expected to match the sign of the NH forced changes (as S15 proposed) and thus may be one reason why the S15 aerosol constraint is a poor predictor of CMIP5 temperature changes.

4. Choice of past aerosol emission estimate?

S15 uses the historical SO_2 emission inventory used in the CMIP5 generation of models (Smith et al. 2011). However, repeating the same analysis using the historical SO_2 emissions based on Smith et al. (2004) (used by many CMIP3 generation models) leads to a much broader range of 2005 aerosol forcings consistent with 1950 observed warming (up to $-1.8\,\mathrm{W\,m^{-2}}$; Fig. 1b). The Smith et al. (2004) emissions are slightly lower in 1950 and higher in the present. If we repeat the analysis including both the Smith et al. (2004) emissions and assuming global linearity (see the discussion above of linear forcing) then this constraint is unable to reject any aerosol forcings up to $-2.1\,\mathrm{W\,m^{-2}}$ (Fig. 1b).

We are not arguing here that the Smith et al. (2004) emission estimates are more plausible than those of Smith et al. (2011). The fundamental point is that if S15 can be said to represent a constraint on aerosol forcing then any real-world application would need to also sample plausible uncertainties in historical SO_2 emission reconstructions. The Smith et al. (2004) emissions sit comfortably inside Smith et al.'s (2011) current estimated uncertainty for most of the time series (Smith et al. 2011, see Fig. S-6 therein). Repeating the S15 energy balance analysis with Smith et al. (2004) illustrates the impact of plausible uncertainty in these reconstructions and shows that more negative real world aerosol forcings are likely to be consistent with the S15 method if S15 is extended to account for this uncertainty.

5. Choice of time period for constraint and energy balance limitations

S15 chose preindustrial to 1950 warming as two dates that avoided large volcanic events and occurred sufficiently

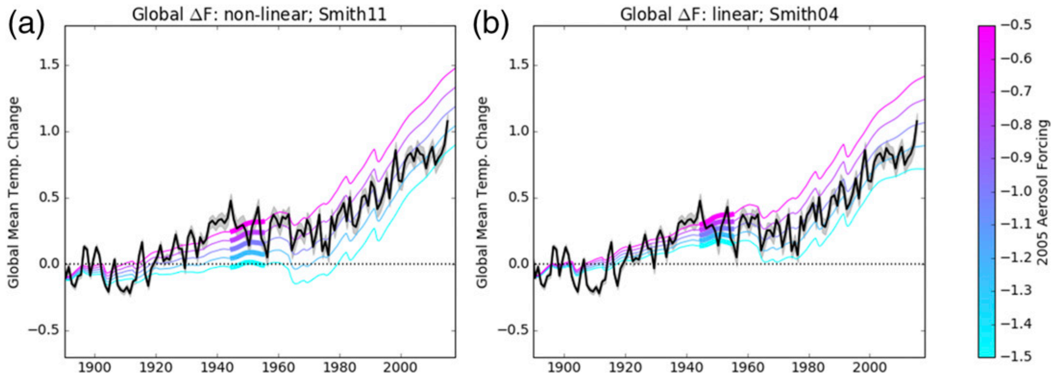

FIG. 2. (a) The time series illustrating the dependence of global mean temperatures when varying on the magnitude of global 2005 aerosol forcing (denoted by the color bar) from an energy balance model/simple climate model [described in Harris et al. (2013)]. The temperatures in the 11 years centered around 1950 are highlighted by the thicker lines. The SO_2 emissions are taken from Smith et al. (2011) and the indirect component of the global forcing is treated as logarithmic to these emissions (as per S15). The other simple model parameters are set at "central" values (climate sensitivity = 3 K; 2 × CO_2 radiative forcing = 3.71 W m^{-2}; land–sea contrast = 1.4; fraction of aerosol forcing treated as linear = 0.22; background natural emissions = 31.3). The mean (black) and uncertainty range (gray) for an observational estimate of historical temperature change from Morice et al. (2012) are shown for comparison. (b) As in (a), but using Smith et al. (2004) SO_2 emissions and treating the global forcing as a linear function of these emissions.

early that any nonlinear forcing response could be accounted for. However, if we were able to account for volcanic forcings (as we do here) and the global (if not regional) aerosol forcing responds linearly with past emissions changes (as in the discussion above of linear forcing) then there are no reasons to choose this over any other period. In the wider context of the twentieth century, the 1950s are unusual in that they tend to be consistent with lower estimates of aerosol forcing. Figure 2a provides an illustration using a simple climate model [documented in Harris et al. (2013)] to translate forcing time series into global temperature changes, while keeping the Smith et al. (2011) emissions and nonlinear indirect aerosol representation used in S15. This shows how varying the aerosol forcing (while fixing other parameters at their standard values) leads to a span of historical temperature changes. When the S15 assumptions are retained, the simple model illustrates why 1950 suggests larger aerosol forcings are less consistent. Lower estimates of aerosol forcing are more consistent in early periods (present-day estimates on the order of −0.5 W m^{-2} are more consistent in the 1950s); however, these tend to overestimate the warming in the later periods (when present-day aerosol forcings in the −1.3 W m^{-2} ballpark are most consistent with 1990s temperatures). The reason why the simple model struggles to reproduce both early and late twentieth-century values is not clear. This could point to outlining combinations of climate sensitivity and heat uptake parameters required for a given aerosol forcing to match the whole record, or (as we go on to show) may instead relate to uncertain assumptions about historical aerosol emissions and how we assume global aerosol forcing to be related to these. This wider twentieth-century context, however, suggests that caution is required before putting too much weight on a particular aerosol forcing period (such as the 1950s) as these forcing may not be reflective of the magnitude of aerosol changes required to reconcile later twentieth-century temperatures changes.

6. Outlook on the potential for requiring net positive 1950 energy balance to constrain the range of real-world aerosol forcing

The S15 constraint is attractive in that it provides a way to reduce present-day aerosol forcing uncertainty using simple and easily understood arguments. The central argument of our comment, however, is that this approach fails to correctly predict a lack of forced trends in any CMIP5 models where aerosol forcing exceeds S15's thresholds. Understanding the cause of this inconsistency is not straightforward, but we have identified several factors that may contribute. Stevens himself acknowledges that sensitivity to emissions and assumptions of linearity are likely to exist:

> "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 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" (p. 4811, S15).

However, from S15 it was not evident what impact these choices would have on the constraint, which is what we have been able to estimate here. By changing either of these two factors [by assuming that the aerosol forcing responds linearly (globally if not locally) or by changing the SO_2 emissions to Smith et al. (2004)] we show that the implied constraint changes from one that rules out aerosol forcing ranges simulated by the larger fraction of aerosol indirect effect capable CMIP5 simulations to a constraint that rules out none.

Concerns over the particular choice of time period aside, globally we see merits in exploring requirements for net forcing to be positive as a potential constraint. However, big questions need to be asked about such an approach that fails to predict the sign of forced temperature changes in CMIP5 models. Kretzschmar et al. (2017) have already used other inferences of CMIP5 historical aerosol forcing to highlight that the logarithmic global relationship of forcing with sulfur emissions used by S15 (and elsewhere) may be one factor behind this failure. In response Stevens and Fiedler (2017) identified behavior in the CMIP5 models, used in Kretzschmar et al. (2017), that questions the plausibility of these models' historic forcing responses. However, it is difficult to assess to what extent the individual modeling errors identified imply that the logarithmic relationship employed in S15 would be a better model. Both Stevens and Fiedler 2017 and S15 put forward data supporting this relationship but these forcing estimates cover only a handful of dates [only two in Stevens and Fiedler (2017) and four in S15], none of which cover the 1900–75 period. This makes it difficult to have confidence in a logarithmic relationship, given that forcing estimates from Kretzschmar et al. (2017) and the new data presented in this comment highlight the potential for more linear global responses. Perhaps the best that can be said is that S15 and these subsequent comments highlight the need for more extensive time-evolving aerosol forcing estimates to better understand the linearity of global forcing (or otherwise).

Moving beyond CMIP5, we have shown that the application of the S15 constraint to the real world will be dependent on the historical sulfur emissions used. If the Smith et al. (2004) emissions are used instead and we assume forcing linearity, then the global S15 approach would not rule out any present-day aerosol forcings up to $-2.1\,\mathrm{W\,m^{-2}}$ (Fig. 1b). Figure 2b helps illustrate why 1950's temperature rise is less effective a constraint in this case. In contrast to Fig. 2a, Fig. 2b uses the Smith et al. (2004) emissions and assumes a globally linear aerosol forcing response to emissions. Consequently, global temperature changes estimated from a simple climate model are more similar up to the 1950s (Fig. 2b; cf. Fig. 2a) and none of the simple climate model estimates fail to capture a warming trend. It is perhaps also worth noting that these assumptions make it easier to identify aerosol forcings that are consistent over the whole historical period. Mechanistically both factors [assuming linearity and using the Smith et al. (2004) emissions] lead to a smaller fraction of the present-day aerosol forcing being realized by 1950. The consequence is that net forcing or temperature change to 1950 is much less effective at discriminating amongst the present-day aerosol forcing estimates.

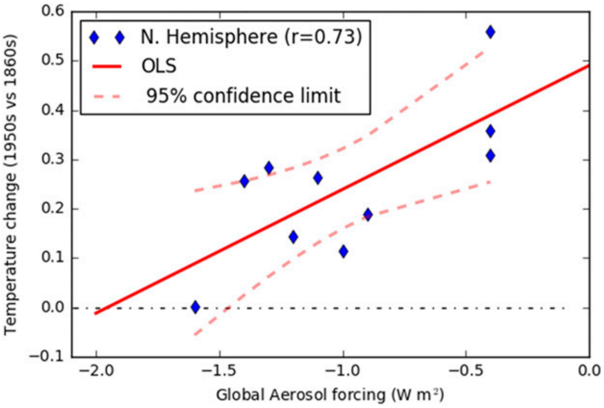

FIG. 3. The relationship between present-day global aerosol forcing and the magnitude of forced (ensemble mean) temperature rise between the 1860s and 1950s from the CMIP5 models used in Table 1. The regressed relationship and 95% confidence limit are shown (red lines) with an intercept of $-2.0\,\mathrm{W\,m^{-2}}$ indicating the threshold above which larger negative aerosol forcings would no longer be expected to be consistent with net positive forced temperature trends in 1950.

We find S15's NH-only approach to be a more questionable constraint. This is the constraint that enabled S15 to rule out present-day aerosol forcings more negative than $-1.0\,\mathrm{W\,m^{-2}}$. By linking the sign of hemispheric forcing to the sign of hemispheric temperature change S15 neglects the role of cross-equatorial energy flux, which we know to be important (see the discussion above of hemispheric constraints). However, CMIP5 models do suggest a fairly strong relationship between the magnitude of global forcing and the forced temperature trend to 1950 ($r = 0.73$), suggesting that it may be worth pursuing. The aerosol forcing explains roughly 50% of the CMIP5 spread in NH forced temperature trends (with presumably differences in other forcings, efficacy of climate sensitivity, and climate model errors also influencing this spread). If we regress the CMIP5 relationship between global present-day forcing and NH forced preindustrial to 1950 temperature change, we get a best estimate constraint of $-2.0\,\mathrm{W\,m^{-2}}$ (Fig. 3). Incidentally, if, as Stevens and Fiedler (2017) argue, there are good reasons to exclude GFDL-CM3, then the implied constraint on aerosol forcings is weaker still. There is substantial uncertainty around this regressed relationship but S15's $-1.0\,\mathrm{W\,m^{-2}}$ constraint is clearly not consistent with this (Fig. 3). This analysis suggests that there may be merit in pursuing hemispheric-only constraints on aerosol forcing but that we do not have the right conceptual framework to do this at present (to account for factors like cross-equatorial energy transfer, for example).

In summary, we conclude that global mean energy balance arguments put forward by S15 imply an overly strong constraint that is inconsistent with what we see in current process-based climate models. Reevaluating the S15 methodology constraint using different, but plausible, assumptions, leads to weaker constraints on the magnitude of present-day aerosol forcing. Any real-world constraint would need to account for historical emission uncertainty and it is difficult to see how this could rule out aerosol forcings less negative than $-2.0\,\mathrm{W\,m^{-2}}$, given the sensitivity to the underlying assumptions shown in this comment.

Acknowledgments. Ben Booth, Glen Harris, and Andy Jones were supported by the Met Office Hadley Centre Climate Programme funded by BEIS and Defra. Matt Hawcroft is supported by the Natural Environment Research Council/Department for International Development via the Future Climates for Africa (FCFA) funded project "Improving Model Processes for African Climate" (IMPALA, NE/ M017265/1). Ken Carslaw was funded by the U.K. Natural Environment Research Council project ACID-PRUF (NE/I020148/1) with support from the Leeds–Met Office Academic Partnership (ASCI project).

REFERENCES

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

Harris, G. R., D. M. Sexton, B. B. B. Booth, M. Collins, and J. M. Murphy, 2013: Probabilistic projections of transient climate change. *Climate Dyn.*, **40**, 2937–2972, https://doi.org/10.1007/s00382-012-1647-y.

Hawcroft, M., J. M. Haywood, M. Collins, A. Jones, A. C. Jones, and G. Stephens, 2017: Southern Ocean albedo, inter-hemispheric energy transports and the double ITCZ: Global impacts of biases in a coupled model. *Climate Dyn.*, **48**, 2279–2295, https://doi.org/10.1007/s00382-016-3205-5.

Haywood, J. M., and Coauthors, 2016: The impact of equilibrating hemispheric albedos on tropical performance in the HadGEM2-ES coupled climate model. *Geophys. Res. Lett.*, **43**, 395–403, https://doi.org/10.1002/2015GL066903.

Kay, J. E., C. Wall, V. Yettella, B. Medeiros, C. Hannay, P. Caldwell, and C. Bitz, 2016: Global climate impacts of fixing the Southern Ocean shortwave radiation bias in the Community Earth System Model (CESM). *J. Climate*, **29**, 4617–4636, https://doi.org/10.1175/JCLI-D-15-0358.1.

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, https://doi.org/10.1175/JCLI-D-16-0668.1.

Loeb, N. G., H. Wang, A. Cheng, S. Kato, J. T. Fasullo, K. M. Xu, and R. P. Allan, 2016: Observational constraints on atmospheric and oceanic cross-equatorial heat transports: Revisiting the precipitation asymmetry problem in climate models. *Climate Dyn.*, **46**, 3239–3257, https://doi.org/10.1007/s00382-015-2766-z.

Mechoso, C. R., and Coauthors, 2016: Can reducing the incoming energy flux over the Southern Ocean in a CGCM improve its simulation of tropical climate? *Geophys. Res. Lett.*, **43**, 11 057–11 063, https://doi.org/10.1002/2016GL071150.

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, https://doi.org/10.1029/2011JD017187.

Smith, S. J., R. Andres, E. Conception, and J. Lurz, 2004: Sulfur dioxide emissions: 1850–2000. Rep. PNNL-14537, 14 pp., www.globalchange.umd.edu/data/publications/PNNL-14537.pdf.

——, 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, https://doi.org/10.5194/acp-11-1101-2011.

Stevens, B., 2015: Rethinking the lower bound on aerosol radiative forcing. *J. Climate*, **28**, 4794–4819, https://doi.org/10.1175/JCLI-D-14-00656.1.

——, and S. Fiedler, 2017: Reply to "Comment on 'Rethinking the lower bound on aerosol radiative forcing.'" *J. Climate*, **30**, 6585–6589, https://doi.org/10.1175/JCLI-D-17-0034.1.

Twomey, S., and P. Squires, 1959: The influence of cloud nucleus population on the microstructure and stability of convective clouds. *Tellus*, **11**, 408–411, https://doi.org/10.3402/tellusa.v11i4.9331.