Article:

Andrews, T and Forster, PM orcid.org/0000-0002-6078-0171 (2020) Energy budget constraints on historical radiative forcing. Nature Climate Change, 10. pp. 313-316. ISSN 1758-678X

https://doi.org/10.1038/s41558-020-0696-1

Energy budget constraints on historical radiative forcing

Timothy Andrews[1] and Piers M. Forster[2]

[1]Met Office Hadley Centre, Exeter, UK.

[2]School of Earth & Environment, University of Leeds, Leeds, UK.

Submitted: 24th May 2019;

Revised: 9th September 2019; 8th November 2019; 6th January 2020.

Radiative forcing is a fundamental quantity for understanding anthropogenic and natural drivers of past and future climate change[1]. Yet significant uncertainty remains in our quantification of radiative forcing and model representation of it[2-4]. Here, we use instrumental measurements of historical global-mean surface temperature change and Earth's total heat uptake, alongside estimates of the Earth's radiative response, to provide a top-down energy budget constraint on historical (1861-1880 to near-present) effective radiative forcing of 2.3 (1.7 to 3.0) [5-95%] Wm^{-2}. This represents a near 40% reduction in the 5-95% uncertainty range assessed by the IPCC Fifth Assessment Report[2]. Although precise estimates of effective radiative forcing in models do not widely exist, our results suggest that the effective radiative forcing may be too small in as many as one third of CMIP5 climate models. Improving model representation of radiative forcing should be a priority for modelling centres. This will reduce uncertainties in climate projections that have stubbornly remained for decades[4,5].

Radiative forcing is a measure of the extent to which anthropogenic activities (such as emissions of carbon dioxide or other greenhouses gases) or natural events (such as volcanic eruptions) alter the flows of energy entering and leaving the Earth's climate system[6]. It is the driver of most other climate changes that follow - whether these be changes in temperature, rainfall or extreme events. It is useful for predicting changes in global-mean surface-air-temperature, ΔT (K), due to its representation in the Earth's energy budget[7], where in response to an effective radiative forcing, F (Wm^{-2}), the change in planetary energy imbalance, ΔN (Wm^{-2}), is given by,

$$\Delta N = F - \lambda \Delta T \qquad (1)$$

where λ ($Wm^{-2}\,K^{-1}$) is the climate feedback parameter and measures the extent to which the climate system radiates heat back out to space in response to ΔT. Note that here we regard F as an 'effective radiative forcing' that is consistent with recent developments in the forcing-feedback paradigm[2,7] (Methods).

Typically Eqn (1) has been used to constrain λ and the Earth's effective climate sensitivity from observed changes in ΔT and ΔN over the historical record[8,9], assuming some estimate of historical F. Here we turn this around[3,10], utilizing new estimates of λ in conjunction with observed estimates of ΔT and ΔN to constrain F. Importantly the generality of Eqn (1) has recently been questioned, since ΔN and λ have been shown to depend on the spatial structure of surface temperature change[11-16], but our approach uniquely accounts for this by using λ estimates that are consistent with the observed geographical patterns of historical climate change, which we call λ_{hist}. λ_{hist} is also estimated independently of F. This is important because previous studies that employed Eqn (1) to estimate radiative forcing have necessarily needed to fit F and λ_{hist} together, leading to difficulty constraining either term accurately from the historical record[17].

Table 1: Comparison of historical (1861-1880 to 2011) effective radiative forcing (ERF) values assessed by IPCC AR5[2] and inferred from CMIP5 climate models[3,29] against our energy budget constraint (1861-1880 to 2005-2015). Uncertainties are 5 to 95%. Also shown are sensitivity tests to assumptions in our methodology (Methods).

	Total forcing	Aerosol forcing
IPCC AR5 assessed range of ERF	**2.2 (1.0 to 3.2)**	**-0.7 (-1.7 to +0.1)**
CMIP5 climate model range of inferred ERF	**1.9 (1.0 to 2.8)**	**-1.2 (-1.7 to -0.7)**
This study's historical energy budget constraint on ERF	**2.3 (1.7 to 3.0)**	**-0.8 (-1.6 to +0.1)**
This study with 1850-1900 baseline	2.3 (1.7 to 3.1)	-0.7 (-1.5 to +0.1)
This study with 2005-2017 present-day	2.4 (1.7 to 3.1)	-0.7 (-1.5 to +0.2)
This study with blended SST and SAT anomalies	2.2 (1.5 to 2.9)	-0.9 (-1.7 to -0.1)
This study with +0.1 Wm^{-2} larger pre-industrial heat uptake	2.2 (1.6 to 2.9)	-0.9 (-1.7 to 0.0)
This study with 50% inflation in λ_{hist} uncertainty	2.3 (1.5 to 3.2)	-0.8 (-1.7 to +0.3)

We define our pre-industrial baseline as 1861-1880 (inclusive)[18] and present-day to be 2005-2015 (inclusive) in order to agree with the time period of recently observed energy flows[19] and readily allow comparisons to Intergovernmental Panel on Climate Change Fifth Assessment Report (IPCC AR5) 'present-day' (year 2011) assessed radiative forcings and their uncertainties[2]. For ΔT we use global-annual-mean temperature anomalies derived from blended sea-surface-temperatures (SST) (over oceans) and surface-air-temperatures (SAT) (over land) with complete global coverage[20] scaled to a SAT anomalies (see Methods). For 2005-2015 minus 1861-1880, this gives $\Delta T = 0.98 \pm 0.20$ [5-95%] K, where the uncertainty accounts for both instrumental errors and internal variability[9,18]. For the Earth's energy imbalance we take the present-day (2005 to 2015) observed value of 0.71 ± 0.10 [5-95%] Wm^{-2} from Johnson et al.[19], derived from in-situ ocean observations with near-global coverage. For the change in imbalance, ΔN, we allow for a pre-industrial value of 0.10 Wm^{-2} following Forster[21], giving $\Delta N = 0.61 \pm 0.35$ [5-95%] Wm^{-2}, where the uncertainty is based on model ensemble spread (see Methods).

For λ_{hist} we use a suite[22] of Atmospheric General Circulation Models (AGCMs) forced with observed (AMIP II boundary conditions[23]) changes in monthly SST and sea-ice variations from 1871 to 2010. AGCMs forced in this way – while keeping all forcing agents such as greenhouse gases and aerosols constant – allows a straightforward way of calculating the radiative response to changes in observed SST and sea-ice that occurred over the historical period from a single linear fit between ΔN and ΔT over the entire historical period[13,14,22] (Methods). When constrained by real world SST and sea-ice, AGCM radiative feedbacks represent the flux anomalies that occur in response to real world historical SST and sea-ice variations, and compare extremely well against observations over a range of surface temperature patterns on interannual (e.g. in response to ENSO) to decadal timescales[14,24-26]. While no single model can be assumed to reproduce λ_{hist} precisely, Andrews et al.[22] give $\lambda_{hist} = 1.74 \pm 0.48$ [5-95%] $Wm^{-2} K^{-1}$ across a model ensemble of eight AGCMs, which we use here for our λ_{hist} and its uncertainty. While this uncertainty is derived from model spread from a limited set of models, it is expected to cover a similar uncertainty to that which would arise from a larger model ensemble, but it does not include structural uncertainties related to model error (Methods).

The principle value of using λ_{hist} derived from AGCMs forced with the observed patterns of SST and sea-ice variations is that it ensures that the resulting λ_{hist} is consistent with the geographical patterns of historical temperature change that gave rise to the non-radiative forcing component of the observed ΔN values (i.e. $\lambda\Delta T$). This is crucial since ΔN and λ have been shown to depend on the spatial structure of surface temperature change[11-16]. Our approach ensures λ_{hist} and ΔN are consistent with the same observed SST pattern, regardless of how the pattern of historical temperature change arose – whether it be affected by internal variability, different timescale responses in the climate system or due to specific impacts of spatiotemporal variations in anthropogenic or natural forcings (Methods). Our approach assumes that applying our method to different realisations of internal variability in historical climate change would result in the same F (Methods). In contrast, assuming a λ derived from any non-observed pattern of climate change, such as from CO_2 forced climate sensitivity experiments, may be inaccurate for use over the historical record due to differences in temperature patterns – and so feedbacks – between the observed historical record and those simulated under long-term CO_2 forced climate change[13,14,16].

To generate our constraint on F we randomly sample (with replacement) 1 million times from the Gaussian distributions of ΔT, ΔN and λ_{hist} to calculate $F = \Delta N + \lambda_{hist}\Delta T$. The resulting values are binned into intervals of 0.02 Wm^{-2} and normalised to produce a probability density function (PDF), excluding values outside ±10 Wm^{-2} (this choice has no impact on our results). The resulting PDF (Figure 1, black line) gives a total forcing (median) of 2.3 (1.7 to 3.0) [5-95%] Wm^{-2} (Table 1) and is very nearly - though not quite - Gaussian. IPCC AR5 assessed various lines of evidence on radiative forcing, including (but not exclusively) (i) line-by-line radiative transfer calculations and laboratory measurements of spectral absorption; (ii) bottom-up process understanding of radiative forcing and rapid 'adjustment' mechanisms; (iii) satellite observations of radiative fluxes and aerosol-cloud properties; and (iv) a range of climate modelling from high resolution cloud-resolving simulations to complex global AGCMs. They assessed the anthropogenic radiative forcing (corrected to an 1861-1880 baseline, see Methods) to be 2.2 (1.0 to 3.2) [5-95%] Wm^{-2}. We have independently arrived at a similar best estimate, derived simply from a top-down energy budget constraint, but with a near 40% reduction in the 5-95% uncertainty range and a substantial reduction in the probability of radiative forcings at the lower end of the IPCC AR5 range (the 5th percentile increases from 1.0 to 1.7 Wm^{-2}) (Table 1 and Figure 1).

How well do climate models represent effective radiative forcing? Radiative forcings are not currently directly calculated in models and so must be inferred indirectly. Forster et al.[3] estimated the historical F of 23 Atmosphere Ocean General Circulation Models (AOGCMs) that participated in CMIP5 using a similar energy budget approach to that used here, but with the limitation of necessarily having to assume a λ diagnosed from long-term CO_2 quadrupling climate sensitivity experiments (λ_{4x}) rather than a more accurate model specific λ_{hist}. This has the potential to bias model estimates of F low if $\lambda_{4x} < \lambda_{hist}$ in AOGCMs (see Methods). They inferred a modelled F of 1.9 (1.0 to 2.8) [5-95%] Wm^{-2}. This falls within the IPCC AR5 5 to 95% uncertainty (Table 1; Figure 1), but compared to our tighter constraint the effective radiative forcing may be too small in as many as one third (8 out of 23) CMIP5 AOGCMs (Table 1; Figure 1), i.e. below our 5th percentile (1.7 Wm^{-2}). However no model has an F above our 95th percentile (3.0 Wm^{-2}).

The radiative effect of changes in anthropogenic aerosol forcing – in particular from aerosol-cloud interactions – has long been identified as the largest source of uncertainty in anthropogenic radiative forcing[2] and some argue[27,28] it is overly negative in models. Our top-down energy budget constraint does not distinguish between individual drivers of climate change. However by assuming a distribution on the non-aerosol forcings (3.06 ± 0.45 [5-95%] Wm^{-2}, Methods) we arrive at a residual aerosol forcing of -0.8 (-1.6 to +0.1) [5-95%] Wm^{-2}. This is in good agreement with the IPCC AR5 assessment (Table 1; Figure 1), providing further confidence in this range. In comparison to nine CMIP5 models assessed by Zelinka et al.[29] (see Methods) nearly all (8 out of 9) models have an aerosol forcing in the lower (strongly negative) half of this distribution (Figure 1; Table 1).

We check the robustness of our energy budget constraint on F to various assumptions and structural uncertainties with sensitivity tests (Methods and Table 1). For example using different time-periods for the pre-industrial or present-day increases F by no more than 0.1 Wm^{-2} (Methods and Table 1). Allowing for a larger pre-industrial energy imbalance, for example up to the maximum estimate of 0.2 Wm^{-2} suggested by Gregory et al.[8], simply shifts our forcing distributions by a corresponding amount (i.e. -0.1 Wm^{-2} in this case, Methods and Table 1). Inflating our uncertainty in λ_{hist} by 50% (to test the sensitivity of our results to unknown structural model errors) changes our constraint by no more than ±0.2 Wm^{-2} (Methods and Table 1) and even against this increased 5-95% confidence interval nearly one third (7 out of 23) of CMIP5 models still have estimated forcings below this range.

Here we have provided a top-down energy budget constraint on effective radiative forcing since pre-industrial times. The novel step is to recognise that instrumental measurements of historical global-mean surface temperature change and Earth's total heat uptake can be combined with reconstructions of the Earth's radiative response – derived from AGCMs forced with observed spatial patterns of historical SST and sea-ice change – to constrain historical effective radiative forcing. AGCMs constrained in this way provide a useful constraint on the Earth's radiative feedback over the historical period and compare extremely well against observations over a range of surface temperature patterns on interannual to decadal timescales[14,24-26]. However the use of models inevitably leaves the potential for structural errors in our constraint, such as unknown model bias in λ_{hist}. Still, our energy budget constraint suggests that as many as one third of CMIP5 climate models may have historical radiative forcings that are too small, and nearly all have an aerosol forcing in the lower (strongly negative) half of our aerosol forcing constraint. It has been suggested[27] that aerosol forcing is overly negative in models, perhaps due to missing liquid water path 'buffering' processes seen in comparisons to observations[28,30]. Alternatively a potentially weak total forcing could additionally arise from positive greenhouse gas forcings that are too weak in models[31] or other uncertainties such as a strong negative land-use forcing seen in some Earth system models[32].

If historical forcing is too weak in models then they may struggle to accurately simulate historical temperature trends[3,33] or have compensating inaccuracies in their historical heat uptake or effective climate sensitivities[34]. Inaccuracies in coupled atmosphere-ocean simulations of λ_{hist} might be expected given AOGCMs struggle to simulate recent decadal temperature trends in the Pacific[14,16], but does not necessarily imply their long-term sensitivity to CO_2 is inaccurate[22]. Improved simulation of historical effective radiative forcing benchmarked against our energy budget constraint would help reduce one aspect of uncertainty in simulated historical temperature trends; improving the detection and attribution of climate change[35] and better illuminating model diversity in transient heat uptake and climate sensitivity processes that could be more readily evaluated against observed temperature trends. Since effective radiative forcings are not routinely diagnosed directly in models various methods have been developed to infer them, and some assumptions – such as assuming invariance in feedbacks[3] – may lead to inaccuracies or even biases in the inferred model forcings, such as those reported for CMIP5 models. In the future, participation by modelling centres in the Radiative Forcing Model Intercomparison Project[5] (RFMIP) will relieve this issue, since historical effective radiative forcing will be directly diagnosed from models. This will reveal model uncertainty in radiative forcing in a more accurate and comprehensive way, enabling a rigorous comparison to process understanding, observations and top-down energy budget constraints as presented here. Significant uncertainty in the model representation of radiative forcing remains[3-5] and our energy budget constraint will enable model improvements aimed at reducing this spread, thus reducing uncertainty in projections of 21st century radiative forcing and climate change[4].

References

[1] Shine, K.P., and P.M.Forster, 1999: The effect of human activity on radiative forcing of climate change: a review of recent developments. *Glo. Plan. Ch.*, 20, 4, doi:10.1016/S0921-8181(99)00017-X.

[2] Myhre, G., et al., 2013: Anthropogenic and natural radiative forcing. In Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. T.F. Stocker, et al., Eds. Cambridge University Press, pp 659-740.

[3] Forster, P. M., T. Andrews, P. Good, J. M. Gregory, L. S. Jackson, and M. Zelinka, 2013: Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models. *J. Geophys. Res.*, 118, 1139–1150, doi: 10.1002/jgrd.50174.

[4] Soden, B.J., W.D. Collins and D.R. Feldman, 2018: Reducing uncertainties in climate models. *Science*, 326-327, doi: 10.1126/science.aau1864.

[5] Pincus, R., P.M. Forster and B. Stevens, 2016: The Radiative Forcing Model Intercomparison Project (RFMIP): experimental protocol for CMIP6, *Geosci. Model Dev.*, 9, 3447-3460, doi: 10.5194/gmd-9-3447-2016.

[6] National Research Council, 2005: Radiative Forcing of Climate Change. *National Academies Press*, 207 pp.

[7] Sherwood, S. C., et al., 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.

[8] Gregory, J.M., R.J. Stouffer, S.C. Raper, P.A. Stott, and N.A. Rayner, 2002: An Observationally Based Estimate of the Climate Sensitivity. *J. Climate*, 15, 3117–3121.

[9] Otto., A., et al., 2013: Energy budget constraints on climate response. *Nature Geoscience*. 6, 415-416, doi:10.1038/ngeo1836.

[10] Forster, P.M. and K.E. Taylor, 2006: Climate Forcings and Climate Sensitivities Diagnosed from Coupled Climate Model Integrations. *J. Climate*, 19, 6181–6194, doi: 10.1175/JCLI3974.1.

[11] Andrews, T., J.M. Gregory and M.J. Webb, 2015: The dependence of radiative forcing and feedback on evolving patterns of surface temperature change in climate models. *J. Climate*, 28, 1630-1648, doi:10.1175/JCLI-D-14-00545.1.

[12] Andrews, T., and M.J. Webb, 2018: The dependence of global cloud and lapse rate feedbacks on the spatial structure of tropical pacific warming. J. Climate, 31, 641–654, doi:10.1175/JCLI-D-17-0087.

[13] Gregory, J.M., and T. Andrews, 2016: Variation in climate sensitivity and feedback parameters during the historical period. *Geophys. Res. Lett.*, 43, 3911–3920, doi:10.1002/2016GL068406.

[14] Zhou., C., M.D. Zelinka and S.A. Klein, 2016: Impact of decadal cloud variations on the Earth's energy budget. *Nature Geoscience*, 9, 871-874, doi:10.1038/ngeo2828.

[15] Xie, S.-P., Y. Kosaka, and Y.M. Okumura, 2016: Distinct energy budgets for anthropogenic and natural changes during global warming hiatus. *Nature Geoscience*, 9(1), 29–33. https://doi.org/10.1038/ngeo2581.

[16] Dong, Y., C. Proistosescu, K.C. Armour, and D.S. Battisti, 2019: Attributing Historical and Future Evolution of Radiative Feedbacks to Regional Warming Patterns using a Green's Function Approach: The Preeminence of the Western Pacific. *J. Climate*, 32, 5471–5491, doi: 10.1175/JCLI-D-18-0843.1

[17] Forest, C.E., et al., 2002: Quantifying uncertainties in climate system properties with the use of recent climate observations. *Science*. 295: 113-117. doi: 10.1126/science.1064419.

[18] Richardson, M., K. Cowtan, E. Hawkinsand M.B. Stolpe, 2016: Reconciled climate response estimates from climate models and the energy budget of Earth. *Nature Climate Change*, 6, pages 931–935.

[19] Johnson, G.C, J.M. Lyman and N.G. Loeb, 2016: Improving estimates of Earth's energy imbalance. *Nature Climate Change*. 6, 639–640.

[20] Cowtan, K. and Way, R. G. (2014), Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. Q.J.R. Meteorol. Soc., 140: 1935-1944. doi:10.1002/qj.2297.

[21] Forster, P.M., 2016: Inference of Climate Sensitivity from Analysis of Earth's Energy Budget. *Annual Review of Earth and Planetary Sciences*, 44:85-106, doi: 10.1146/annurev-earth-060614-105156.

[22] Andrews, T., et al., 2018: Accounting for changing temperature patterns increases historical estimates of climate sensitivity. *Geophysical Research Letters*, 45, 8490– 8499, doi: 10.1029/2018GL078887.

[23] Hurrell, J., et al., 2008: A New Sea Surface Temperature and Sea Ice Boundary Dataset for the Community Atmosphere Model. *J. Climate*, 21, 5145–5153, doi: 10.1175/2008JCLI2292.1.

[24] Allan, R. P., C. Liu, N. G. Loeb, M. D. Palmer, M. Roberts, D. Smith and P.-L. Vidale, 2014: Changes in global net radiative imbalance 1985-2012, *Geophys. Res. Lett.*, doi: 10.1002/2014GL060962.

[25] Smith, D. M., Allan, R. P., Coward, A. C., Eade, R., Hyder, P., Liu, C., Loeb, N. G., Palmer, M. D., Roberts, C. D. and Scaife, A. A., 2015: Earth's energy imbalance since 1960 in observations and CMIP5 models. *Geophys. Res. Lett.*, 42: 1205– 1213. doi: 10.1002/2014GL062669.

[26] Ceppi, P., and J.M. Gregory, 2017: Relationship of tropospheric stability to climate sensitivity and Earth’s observed radiation budget. *PNAS*, 114(50), doi:10.1073/pnas.1714308114.

[27] Stevens, B., 2015: Rethinking the Lower Bound on Aerosol Radiative Forcing. *J. Climate*, 28, 4794–4819, doi:10.1175/JCLI-D-14-00656.1.

[28] Sato, Y., et al., 2018: Aerosol effects on cloud water amounts were successfully simulated by a global cloud-system resolving model, *Nat. Commun*. 9(1), 985, doi:10.1038/s41467-018-03379-6.

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

[30] Toll, V., Christensen, M., Gassó, S., & Bellouin, N., 2017: Volcano and ship tracks indicate excessive aerosol-induced cloud water increases in a climate model. *Geophysical Research Letters*, 44, 12492–12500. doi: 10.1002/2017GL075280.

[31] Etminan, M., G. Myhre, E.J. Highwood and K.P Shine, 2016: Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing. *Geophys. Res. Lett.* 43, 12, 614-12, 623, doi: 10.1002/2016GL071930.

[32] Andrews, T., R.A. Betts, B.B.B. Booth, C.D. Jones and G.S. Jones, 2016: Effective radiative forcing from historical land-use change. *Climate Dynamics*, doi:10.1007/s00382-016-3280-7.

[33] Golaz, J.-C., et al., 2019: The DOE E3SM coupled model version 1: Overview and evaluation at standard resolution. *Journal of Advances in Modeling Earth Systems*, 11, 2089– 2129, doi: 10.1029/2018MS001603.

[34] Kiehl, J. T., 2007: Twentieth century climate model response and climate sensitivity, *Geophys. Res. Lett.*, 34, L22710, doi:10.1029/2007GL031383.

[35] Hegerl, G. and F. Zwiers 2011: Use of models in detection and attribution of climate change. *WIREs Clim Change*, 2: 570-591. doi:10.1002/wcc.121.

Figure Captions

Figure 1: Historical effective radiative forcing (1861-1880 to near-present) derived from an energy budget constraint compared to that assessed by IPCC AR5 and inferred from CMIP5 climate models. (black) Historical effective radiative forcing (1861-1880 to near-present) probability distribution function derived from historical energy budget constraints (Eqn 1). (Blue) the aerosol component. (Red) IPCC AR5 assessed range[2] of historical effective radiative forcing and (gray) inferred CMIP5 climate model effective radiative forcings[3,29]. Gray dots represent individual CMIP5 models. Lines and bars represent the best estimate and 5-95% confidence interval.

Methods

Effective radiative forcing (ERF): Radiative forcing can be measured in different ways and ERF is now a widely adopted definition[2,5,7]. ERF not only includes the instantaneous radiative effect of a change in forcing agent (such as the immediate reduction in outgoing radiation to space in response to increased greenhouse gases) but also any other responses to forcing that are not mediated by the global-mean surface-air-temperature change, ΔT. Examples of such 'adjustments' include aerosol-cloud interactions and stratospheric temperature adjustment. Sherwood et al.[7] provide a pedagogical review of these recent developments. We regard our constraint on F to be appropriate to ERF since we use a feedback parameter (λ) consistent with the ERF framework, so that $\lambda\Delta T$ in Eqn (1) accounts for ΔT-mediated component of ΔN.

Global surface-air-temperature data: We use the Cowtan and Way[20] global-annual-mean temperature anomaly dataset derived from blended sea-surface-temperatures (SST) (over oceans) and surface-air-temperatures (SAT) (over land) with complete global coverage (https://www-users.york.ac.uk/~kdc3/papers/coverage2013/had4_krig_annual_v2_0_0.txt; accessed 16th May 2019). Blended SST and SAT datasets will understate the magnitude of historical warming relative to a SAT only definition[18,36] of ΔT as demanded by Eqn (1). We allow for this following Richardson et al.[16], who found that climate models closely agreed on a scaling of ~1.09 from blended historical SST and SAT anomalies to SAT anomalies. For 2005-2015 minus 1861-1880 (inclusive), the Cowtan and Way[20] dataset gives a historical temperature anomaly of 0.90 K, which - after scaling to SAT - gives ΔT = 0.98 ± 0.20 [5-95%] K, where the uncertainty is taken from Richardson et al.[16] and Otto et al.[9] and accounts for both instrumental errors and internal variability.

Planetary energy imbalance data: For the present-day Earth's energy imbalance we use the 2005 to 2015 observed value of 0.71 ± 0.10 [5-95%] Wm^{-2} from Johnson et al.[19], derived from in-situ ocean observations with near-global coverage. Since heat uptake was not observed in the 19th century, climate and energy balance models must be used to infer the pre-industrial energy imbalance[8,9,37,38]. Forster[21] summarise this be around 0.10 Wm^{-2}, which we use in our best estimate, but allow for a large uncertainty which we also take from Forster[21]. They used 25 AOGCM simulations of historical ΔN to arrive at a 5 to 95% uncertainty range of 0.33 to 1.04 Wm^{-2}, which we approximate as a Gaussian of ± 0.35 Wm^{-2}. Hence we use ΔN = 0.61 ± 0.35 [5-95%] Wm^{-2}. The total uncertainty is difficult to quantify formally, but we have allowed for more than most previous studies[9,38]. The use of models unavoidably leaves the potential for structural errors and uncertainties, for example from in the implementation and spin-up of volcanic forcing across models[39]. Though the impact of volcanism is partly minimised here, since Forster[21] used a reference (1860-1880) and final (2000-2012) period that were both volcanically quiet periods. Note also that if the pre-industrial value was to shift, for example from our assumed 0.1 Wm^{-2} to the maximum estimate of 0.2 Wm^{-2} given by Gregory et al.[8], then this simply shifts our forcing distributions by a corresponding amount (0.1 Wm^{-2} in this case, Table 1 and see Sensitivity tests in Methods).

Climate feedback parameter data: For the historical climate feedback parameter, λ_{hist}, we use the mean and 5 to 95% uncertainty (λ_{hist} = 1.74 ± 0.48 [5-95%] Wm^{-2} K^{-1}) reported in Andrews et al.[22] (their λ_{amip} in Table 1), derived from eight AGCMs forced with observed (AMIP II boundary conditions[23]) changes in monthly SST and sea-ice variations from 1871 to 2010. λ_{hist} represents the linear fit of ΔN and ΔT from these experiments, calculated from an ordinary least square regression over the entire time-period[22] (i.e. 1871 to 2010). The uncertainty – arising from differences in atmospheric parameterisations across models – is derived from the model ensemble spread but does not account for structural uncertainties related to model error and any dependence on the underlying SST and sea-ice historical datasets forcing the AGCMs (see Sensitivity tests in the Methods). However the uncertainty range is of similar magnitude to that seen in traditional CO_2 forced climate sensitivity experiments but shifted towards larger λ (smaller climate sensitivity)[22] (the standard deviation in λ is 0.29 Wm^{-2} across 8 models in Andrews et al.[22], compared to 0.31 Wm^{-2} across 23 CMIP5 AOGCMs forced by abrupt-4xCO_2 in Forster et al.[3]) Hence we have little reason to doubt that the small ensemble of models covers a similar uncertainty in feedback that would exist in a larger ensemble of models.

Insensitivity of *F* to the pattern of change and variability: Since λ_{hist} is derived using observed patterns of surface temperature change our constraint on *F* will not depend on what that pattern is or how it arose, by construction, assuming that the AGCMs can faithfully capture the radiative response to it. For example, imagine another 'realisation' of the real world where global-mean ΔT was the same but the spatial pattern of internal variability was different over recent decades, so that the eastern Pacific was warming rather than cooling (with compensating changes elsewhere). Such a scenario would plausibly lead to reduced cloudiness in the eastern Pacific due to the local warming of the marine stratocumulus decks and thus a greater observed ΔN[12,14]. However this hypothetical eastern Pacific warming would not bias our estimate of *F*, since the eastern Pacific warming would also be included in the boundary conditions to the AGCMs from which λ_{hist} was derived. Thus under this scenario the inferred λ_{hist} would also be smaller (since the inferred cloud feedback to the observed pattern of climate change would be different), compensating for the greater ΔN. By forcing AGCMs with the observed patterns of historical climate change this ensures λ_{hist} is appropriate for the observed ΔN.

Put another way, different historical realisations will generate different values of ΔN and λ_{hist} for the same historical *F*, as seen in the MPI-ESM1.1 100-member 'grande ensemble' of historical simulations analysed by Dessler et al.[40]. Our approach is analogous to an inverse of this, retrieving the same *F* for potentially different combinations of ΔN and λ_{hist}. A corollary of this interpretation is that variability in surface temperature patterns affect heat uptake and feedbacks in a way that is anti-correlated. This could be tested by applying our methodology to a grande ensemble of historical simulations, provided a corresponding ensemble of analogous *amip-piForcing*[22] simulations (but using the model ensemble generated SST and sea-ice boundary conditions) was generated to diagnose each ensemble member's λ_{hist}. This approach might provide a useful way to further explore the relationship between regional changes in surface temperature patterns, heat uptake and radiative feedbacks[41-43].

IPCC AR5 radiative forcings: IPCC AR5 assessed[2] the total anthropogenic effective radiative forcing over the industrial era (1750 to 2011) to be 2.3 (1.1 to 3.3) Wm^{-2}. To compare against our values we correct the baseline (i.e. 1750 to 1861-1880) by subtracting 0.09 Wm^{-2} to account for the total anthropogenic forcing over this period (IPCC AR5 Table AII.1.2). Similarly, IPCC AR5 assessed the total aerosol effective radiative forcing to be -0.9 (-1.9 to -0.1) Wm^{-2}, to which we add 0.21 Wm^{-2} to account for the aerosol forcing between 1750 to 1861-1880 (IPCC AR5 Table AII.1.2). We ignore the small component from natural forcings in this comparison.

CMIP5 model radiative forcings: The CMIP5 mean and 5-95% range of 23 CMIP5 models' total historical effective radiative forcings are taken from Forster et al.[3] (their Table 2, 'hist 2003' column). An offset of 0.2 Wm^{-2} has been applied, since Forster et al.[3] calculated the CMIP5 historical forcings at year 2003 rather than 2011. The 0.2 Wm^{-2} accounts (primarily) for the growth of CO_2 between 2003 and 2011, derived from the recent decadal trend in forcing of 0.25 Wm^{-2} dec^{-1} given by Myhre et al.[2] (i.e. adjusting to 2011 gives an offset of 8.0*0.25/10.0=0.2 Wm^{-2}). Forster et al.[3] estimated the model historical forcings by first diagnosing a feedback parameter from long-term (150yrs) abrupt-4xCO_2 experiments (λ_{4x}), and assuming it applied equally to the models' historical simulations (λ_{hist}). If $\lambda_{4x} < \lambda_{hist}$ (i.e. climate sensitivity larger in an AOGCMs abrupt-4xCO_2 simulations compared to its historical simulation[44]) then this could bias low the model forcing estimates given in Forster et al.[3] and so hamper the comparison to our energy budget constraint. However they showed their estimate of model radiative forcing was largely robust to this in a single model. Gregory et al.[45] estimated the CMIP5 AOGCM-mean λ_{hist} to be 1.27 $Wm^{-2}\ K^{-1}$, which is slightly larger than the mean λ_{4x} of CMIP5 AOGCMs given by Forster et al.[3] (λ_{4x}=1.13 $Wm^{-2}\ K^{-1}$). If this difference is representative of the 'pattern effect' between historical and abrupt-4xCO_2 simulations in AOGCMs it suggests the forcing estimates of Forster et al.[3] may be biased low by ~0.14 Wm^{-2} in the multi-model mean, given their AOGCM-mean historical dT of ~1.0 K. Future participation by modelling centres in RFMIP[5] will address this bias uncertainty across the model ensemble in a more definitive way.

The CMIP5 mean and 5-95% range of 9 CMIP5 models' aerosol effective radiative forcings are taken directly from Zelinka et al.[29] (their Table 1). While Zelinka et al.[29] assessed the model aerosol radiative forcing at year 2000, the total global-mean aerosol effective radiative forcing has changed little[2] between then and 2011. The model aerosol forcing estimates have been derived from fixed-SST experiments, and so do not suffer from the uncertainty arising from assuming an invariant feedback parameter.

Non-aerosol radiative forcings: The non-aerosol radiative forcing distribution, calculated over 2005-2015 minus 1861-1880, is derived from the forcing dataset described in Dessler and Forster[46]. This is an updated and improved version of the IPCC AR5 forcing and its uncertainty. For example the radiative forcing formula for CO_2, N_2O and CH_4 has been updated according to Etminan et al.[31]. See Dessler and Forster[46] for details.

Sensitivity Tests: We check the robustness of our energy budget constraint on *F* to various assumptions and structural uncertainties with sensitivity tests. (1) The period regarded as 'pre-industrial' is often ill-defined[47,48]. We repeat our calculations using an 1850-1900 pre-industrial baseline instead, which gives ΔT = 1.00 ± 0.20 [5-95%] K. (2) Updating our definition of 'present-day' to include the most recent data (13 year average, 2005-2017, centred on 2011) gives ΔT = 1.02 ± 0.20 [5-95%] K. In either case, the slight upward revision of ΔT relative to the reference calculation increases the radiative forcing by no more than 0.1 Wm^{-2} (Table 1). (3) If we ignore the 1.09 scaling from Richardson et al.[18] to adjust the blended SST-SAT temperature dataset[20] to SAT, then ΔT = 0.90 ± 0.20 [5-95%] K, but the forcing and its uncertainty is reduced by no more than 0.2 Wm^{-2} (Table 1). (4) If we allow for a larger pre-industrial energy imbalance than the 0.1 Wm^{-2} assumed here, for example up to the maximum estimate of 0.2 Wm^{-2} suggested by Gregory et al.[8], then ΔN = 0.51 ± 0.35 [5-95%] Wm^{-2}, but this simply shifts our forcing distributions by a corresponding amount (-0.1 Wm^{-2}, Table 1). There are other structural uncertainties harder to quantify. For example (5) the fidelity of our distribution on λ_{hist} will somewhat depend on the underlying SST and sea-ice boundary conditions that forced the AGCMs from which λ_{hist} was derived. Andrews et al.[22] showed that in two AGCMs λ_{hist} was largely insensitive to two different SST datasets, but did depend on assumptions on the pre-industrial Antarctic sea-ice fractions that are extremely hard to constrain from observations[49]. For example HadISST2.1[49] sea-ice trends result in southern hemisphere radiative feedbacks that are difficult to reconcile with physical understanding of surface albedo feedbacks and those found in AOGCMs[22,50]. We therefore retain the λ_{hist} distribution from Andrews et al.[22] based on the AMIP II SST and sea-ice boundary conditions[23] as our best estimate. However even if we inflate our uncertainty in λ_{hist} by as much as 50%, our constraints on radiative forcing change by no more than ±0.2 Wm^{-2} (Table 1) and even against this increased 5-95% confidence interval nearly one third (7 out of 23) of CMIP5 models still have forcings below this range.

References (Methods only)

[36] Cowtan, K., et al., 2015: Robust comparison of climate models with observations using blended land air and ocean sea surface temperatures. *Geophys. Res. Lett.*, 42, 6526– 6534, doi: 10.1002/2015GL064888.

[37] Armour, K.C., 2017: Energy budget constraints on climate sensitivity in light of inconstant climate feedbacks, *Nature Climate Change*, 7, 331-335, doi: 10.1038/nclimate3278.

[38] Lewis, N. and J. Curry, 2018: The Impact of Recent Forcing and Ocean Heat Uptake Data on Estimates of Climate Sensitivity. J. Climate, 31, 6051–6071, doi: 10.1175/JCLI-D-17-0667.1.

[39] Gregory, J.M, 2010: Long-term effect of volcanic forcing on ocean heat content, *Geophys. Res. Lett.*, 37, L22701, doi:10.1029/2010GL045507.

[40] Dessler, A., T. Mauritsen and B.J. Stevens, 2018: The influence of internal variability on Earth's energy balance framework and implications for estimating climate sensitivity. *Atmos. Chem. Phys.*, 18, 5147–5155, doi:10.5194/acp-18-5147-2018.

[41] Kang, S.M. and S. 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.

[42] Rose, B.E.J., Armour, K.C., Battisti, D.S., Feldl, N., and Koll, D.D.B., 2014: The dependence of transient climate sensitivity and radiative feedbacks on the spatial pattern of ocean heat uptake, *Geophys. Res. Lett.*, 41, 1071– 1078, doi:10.1002/2013GL058955.

[43] Rugenstein, M.A.A., Caldeira, K., and Knutti, R., 2016: Dependence of global radiative feedbacks on evolving patterns of surface heat fluxes, *Geophys. Res. Lett.*, 43, 9877–9885, doi: 10.1002/2016GL070907.

[44] Marvel, K., R. Pincus, G.A. Schmidt and R.L. Miller, 2018: Internal variability and disequilibrium confound estimates of climate sensitivity from observations. *Geophysical Research Letters*, 45, 1595-1601, doi:10.1002/2017GL076468.

[45] Gregory, J.M., T. Andrews, P. Ceppi, T. Mauritsen and M.J. Webb, 2019: How accurately can the climate sensitivity to CO2 be estimated from historical climate change? *Climate Dynamics*, 1-29, doi:10.1007/s00382-019-04991-y

[46] Dessler, A. E., and PM. Forster, 2018: An estimate of equilibrium climate sensitivity from interannual variability. *JGR- Atmos.*, 123, 8634– 8645, doi: 10.1029/2018JD028481.

[47] Schurer, A.P., et al., 2017: Importance of the pre-industrial baseline for likelihood of exceeding Paris goals. *Nature Climate Change*, 7, 563–567, doi: 10.1038/nclimate3345.

[48] Hawkins, et al. , 2017: Estimating Changes in Global Temperature since the Preindustrial Period. *Bull. Amer. Meteor. Soc.*, 98, 1841–1856, doi: 10.1175/BAMS-D-16-0007.1.

[49] Titchner, H. A., and N. A. Rayner (2014), The Met Office Hadley Centre sea ice and sea surface temperature data set, version 2: 1. Sea ice concentrations, *J. Geophys. Res. Atmos.*, 119, 2864–2889, doi:10.1002/2013JD020316.

[50] Schneider, A., M. Flanner, and J. Perket, 2018: Multidecadal variability in surface albedo feedback across CMIP5 models. *Geophysical Research Letters*, 45, 1972– 1980, doi: 10.1002/2017GL076293.

[51] Andrews, Timothy, 2020: Release of Andrews and Forster (2020, NCC) Code (Version v1.0.0). Zenodo. http://doi.org/10.5281/zenodo.3600258.

Acknowledgements

TA was supported by the Met Office Hadley Centre Climate Programme funded by BEIS and Defra. PMF was supported by NERC SMURPHS project NE/N006054/1. Both TA and PMF were supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 820829 (CONSTRAIN project). We thank Mark Ringer and Jonathan Gregory for useful discussions. We are also extremely grateful for positive and constructive reviews from two anonymous reviewers that improved the clarity of the manuscript.

Corresponding author

Timothy Andrews, Met Office Hadley Centre, FitzRoy Road, Exeter, EX1 3PB, UK.

Email: timothy.andrews@metoffice.gov.uk

Data availability

The Cowtan and Way[20] global-annual-mean temperature anomaly dataset was accessed (16th May 2019) via https://www-users.york.ac.uk/~kdc3/papers/coverage2013/had4_krig_annual_v2_0_0.txt. All other data supporting the findings of this study are provided and/or referenced within this paper.

Code availability

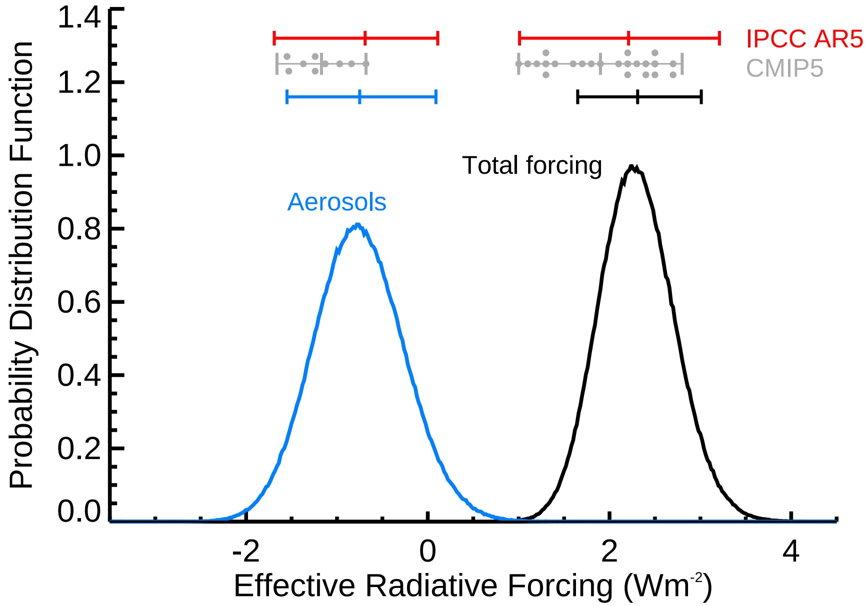
The code used to produce the energy budget constraint and Figure 1 is available via the following online repository: https://github.com/timothyandrews/historical-radiative-forcing[51]

Author contributions

TA and PMF conceived the study. TA performed the analysis and wrote the manuscript, with comments from PMF.

Competing interests

The authors declare no competing interests.



IPCC AR5
CMIP5
Total forcing
Aerosols
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Probability Distribution Function
-2
0
2
4
Effective Radiative Forcing (Wm-2)