Review in Advance first posted online on April 27, 2016. (Changes may still occur before final publication online and in print.)

Inference of Climate Sensitivity from Analysis of Earth's Energy Budget

Piers M. Forster

School of Earth and Environment, University of Leeds, Leeds LS2 9JT, United Kingdom; email: p.m.forster@leeds.ac.uk

Annu. Rev. Earth Planet. Sci. 2016. 44:85–106

The *Annual Review of Earth and Planetary Sciences* is online at earth.annualreviews.org

This article's doi: 10.1146/annurev-earth-060614-105156

Keywords

transient climate response, ocean heat uptake, radiative forcing, aerosol, simple climate models

Abstract

Recent attempts to diagnose equilibrium climate sensitivity (ECS) from changes in Earth's energy budget point toward values at the low end of the Intergovernmental Panel on Climate Change Fifth Assessment Report (AR5)'s likely range (1.5–4.5 K). These studies employ observations but still require an element of modeling to infer ECS. Their diagnosed effective ECS over the historical period of around 2 K holds up to scrutiny, but there is tentative evidence that this underestimates the true ECS from a doubling of carbon dioxide. Different choices of energy imbalance data explain most of the difference between published best estimates, and effective radiative forcing dominates the overall uncertainty. For decadal analyses the largest source of uncertainty comes from a poor understanding of the relationship between ECS and decadal feedback. Considerable progress could be made by diagnosing effective radiative forcing in models.

1. INTRODUCTION

The equilibrium climate sensitivity (ECS)—defined as the globally averaged surface temperature change at equilibrium for a doubling of atmospheric CO_2 concentration—has taken on almost a mythical status in climate science as the uncertainty that will not go away. From its introduction and first assessment by Charney et al. (1979) until the latest Intergovernmental Panel on Climate Change (IPCC) report, the Fifth Assessment Report (AR5) (Stocker et al. 2013), its range has hovered around 1.5–4.5 K. This temperature range represents the uncertainty in global mean warming for a sustained doubling of background carbon dioxide levels.

The seeming intransigence of the uncertainty range belies huge improvements in our understanding and quantification of many aspects of climate, including its sensitivity. It is too simplistic to compare the 1979 range to the 2013 range and say that we are no better off. The 1979 Charney range was a subjective judgment based on two basic global climate models. In contrast, the 2013 AR5 range was based on an assessment of multiple lines of evidence, including comparing the latest climate models to observations, understanding historical records and the paleoclimate record, and investigating changes in Earth's energy budget. Additionally, the science community has realized that a factor-of-three uncertainty in ECS does not necessarily imply a factor-of-three uncertainty in projected warming. A related quantity, the transient climate response (TCR), is much more important for defining warming during the coming century. TCR is defined as how much the world would warm after 70 years of continually increasing CO_2 levels (at 1% per year) and has a smaller uncertainty range [1.0–2.5°C (Stocker et al. 2013)]. As well as the uncertainty in TCR, future climate is affected by uncertainties in projected emissions and feedbacks with the Earth system (such as methane release from Arctic tundra). Further, the impacts of climate are locally realized, and future regional rainfall changes (for instance) are not uniquely determined by global temperature change or ECS (Andrews et al. 2010). Nevertheless, ECS remains important to constrain, both for its status as a talisman of climate science and for fundamentally understanding how Earth's temperature responds to perturbations in its energy budget.

Compared to the IPCC's Fourth Assessment Report (AR4) (Solomon et al. 2007), AR5 reduced its lower bound from 2°C to 1.5°C and did not make a best estimate. This was a result of publications deriving ECS from the instrumental record (recent historical temperatures and top-of-atmosphere energy change) tending to report a lower ECS best estimate than did publications using other lines of evidence (see box 12.2, figure 2, of Collins et al. 2013).

Since publication of AR5 it appears that the instrumental-based approaches have continued to diverge from the approaches that compare model climatologies to observation. The purpose of this review article is to critically examine these energy budget approaches and search for possible causes of discrepancy from the different lines of evidence. Section 2 introduces the theoretical basis; Section 3 critically examines past studies, concentrating on those published since AR5; Section 4 examines uncertainty in deriving ECS from multidecadal trends; Section 5 examines shorter-timescale approaches; and Section 6 points to the way forward.

2. THE THEORETICAL BASIS

The time-dependent energy imbalance, N, of Earth (taken as positive downward) can be split into forcing, response, and noise terms such that (e.g., Forster & Gregory 2006, Gregory et al. 2002)

$$N = F - \alpha T + \gamma. \tag{1}$$

F is the radiative forcing; T is the globally averaged surface temperature change; γ represents a noise term whereby N could be affected by variability unrelated to T (discussed in Section 5); and

α is the climate feedback parameter, which gives the dependency of N on T, such that $\alpha = -\frac{\delta N}{\delta T}$. This formulation is a linear approximation to the global energy budget and remains valid provided T is small.

Fourier (1827) was the first to associate planetary temperature with its top-of-atmosphere energy budget. Stefan (1879) used existing laboratory measurements to deduce the relationship between the temperature of a body and its emitted radiation. This relationship became known as the Stefan-Boltzmann law, one of the underpinning laws of physics. Early authors realized that Earth's effective emission temperature and its surface temperature were not the same but could be straightforwardly connected through the gray-body approximation. N could therefore also be written as $N = \mathrm{ASR} - \varepsilon\sigma T_s^4$, where ASR is the absorbed solar radiation, ε the assumed gray-body emissivity of Earth, and T_s the absolute surface temperature. This can be differentiated to give a value for α, assuming Earth has a black-body response, such that $\alpha_{\mathrm{BB}} = 4\varepsilon T_s^3$.

Arrhenius (1896), building on the work of Tyndall (1861), already knew about the role of water vapor and cloud feedbacks modifying the black-body response. Feedbacks can be included in the simple model by adding a temperature dependence in the emissivity and/or ASR terms. The first climate modeling papers included water vapor feedback but typically assumed fixed clouds (e.g., Manabe & Wetherald 1967). A more realistic value of α combines the black-body response from the Stefan-Boltzmann law with that due to climate feedbacks such as water vapor change, lapse rate change, surface reflectance change, and cloud changes: $\alpha = \alpha_{\mathrm{BB}} + \alpha_{\mathrm{w.v.}} + \alpha_{\mathrm{lapse}} \ldots$.

Once the radiative forcing is determined for a doubling of CO_2 ($F_{2\times CO_2}$, approximately 3.7 Wm^{-2}), it then becomes possible to derive ECS by solving Equation 1 for T at equilibrium ($N = 0$). An effective ECS can also be determined for nonequilibrium conditions:

$$\mathrm{ECS} = \frac{F_{2\times CO_2}}{\alpha} = F_{2\times CO_2}\frac{T}{F - N}. \tag{2}$$

It is in making the connection between an α diagnosed in a particular situation to an α or ECS applicable for $2 \times CO_2$ equilibrium conditions that a major fallacy emerges: It is possible to use these sets of equations with observations of N, F, and T to make an estimate of ECS that does not depend on a climate model. Gregory et al. (2002) were the first to employ observations of N directly to make such an "observationally constrained" estimate of ECS. Forster & Gregory (2006, p. 39) overstated the benefits of such an approach by claiming, "Importantly, the [ECS] estimate is completely independent of climate model results." As Equation 1 derives directly from conservation of energy, the Forster & Gregory (2006) claim would appear valid. But it in fact makes the assumption that the α derived from a particular observational period is the same as the α applicable under long-term climate change. Another way of stating this assumption is saying that the effective climate sensitivity (the apparent ECS diagnosed from a specific α) is the same as the true ECS. Uncertainties around the derivation of ECS from an energy budget approach can be attributed to two causes: the model used to translate α into an ECS estimate and the quality of the observation-based data sets.

3. SUMMARY OF PAST RESULTS

The first attempts to constrain ECS from the historical record implicitly assumed a version of Equation 1 and employed some form of climate model to link the variables (Andronova & Schlesinger 2001, Wigley et al. 1997). These methods have continued to be used, most noticeably when linked with a detection and attribution approach (Frame et al. 2005, Lewis 2015). Ocean temperatures are combined with other observations, such as hemispheric temperature differences, to help constrain aspects of the model response (e.g., Aldrin et al. 2012, Knutti et al. 2002). Studies

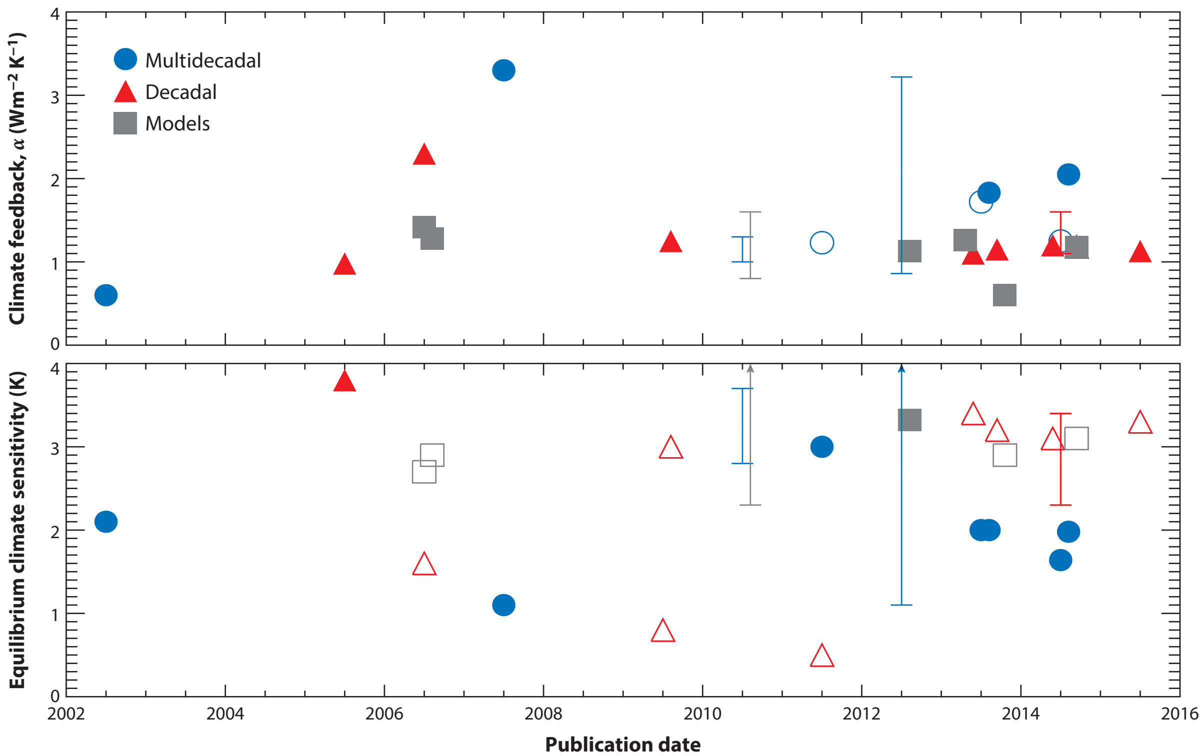


Figure 1

Published best estimate values of the climate feedback parameter, α, and effective equilibrium climate sensitivity by study publication year. These are categorized into multidecadal analyses (*blue circles*), decadal analyses (*red triangles*), and model analyses (*gray squares*). Filled symbols correspond to bold entries in **Table 1**. Four studies quote a range of best estimates illustrated by the error bars (see **Table 1** for details).

have also employed more direct measurements of N, either from ocean heat content (e.g., Gregory et al. 2002, Otto et al. 2013, Roe & Armour 2011, Skeie et al. 2014) or from satellite observations (e.g., Dessler 2013, Forster & Gregory 2006, Lindzen & Choi 2011, Murphy et al. 2009). **Table 1** and **Figure 1** present summaries of the works that have employed variants of Equation 1 to derive α and/or ECS. Studies have been categorized as long-timescale (multidecadal) analyses, short-timescale (decadal) analyses, and climate model analyses. Studies have either employed Equation 1 or 2 directly or used a mixed layer model approach to solve Equation 1, whereby N simply heats a mixed layer ocean of fixed heat capacity, giving the climate a fixed timescale of response (Bengtsson & Schwartz 2013, Schwartz 2007). Other studies have quantified individual feedback terms (Dessler 2013, Soden & Held 2006). Some studies either deliberately or inadvertently omit the forcing term from Equation 1 (Chung et al. 2010; Lindzen & Choi 2009, 2011).

The role of Bayesian uncertainty analysis and prior distributions in making a probabilistic estimate of ECS has been extensively discussed (e.g., Aldrin et al. 2012, Bindoff et al. 2013, Forest et al. 2006, Lewis 2013). A particular online debate surrounded the modification of the prior, which adjusted the Forster & Gregory (2006) estimate of ECS when quoted in IPCC AR4 (Hegerl et al. 2007); see, for example, **http://judithcurry.com/2011/07/05/the-ipccs-alteration-of-forster-gregorys-model-independent-climate-sensitivity-results/**. These choices are important for gauging realistic uncertainty ranges but are not considered further in this review. Rather, this

Table 1 Estimates of α and/or corresponding ECS derived from energy budget changes

Study	Year range	Energy imbalance (N)	Model	α (Wm^{-2} K^{-1})	ECS (K)	Notes
Long-timescale (multidecadal) analyses						
Gregory et al. 2002	1957–1994 cf. 1861–1900	1957–1994 estimate based on 0–700 m OHC from Levitus et al. (2012); $N = 0.16$ Wm^{-2} assumed for base years	Equation 2	**0.6**	**2.1**	Forcing from various sources; α computed from medians in paper; ECS is modal value
Schwartz 2007	1880–2004	1956–2002 estimate based on OHC from Levitus et al. (2012), including deep ocean and other components	Mixed layer model	**3.3**	**1.1**	Forcing not used; time constant and heat capacity of ocean used to determine sensitivity
Lin et al. 2010	1880–2005	Present-day $N =$ 0.85 Wm^{-2} from model data	Mixed layer model	**1.0 to 1.3**	**2.8 to 3.7**	Range depends on ocean heat uptake rates; forcings principally from GISS (Hansen et al. 2007)
Roe & Armour 2011	Preindustrial (not specified)–present	Present-day estimate made in paper is $N = 0.74$ Wm^{-2}	Equation 2	1.23	**3**	Forster et al. (2007) forcings and Solomon et al. (2007) temperatures
Schwartz 2012	1965–2009	1950–2010 various OHC data used, including deep ocean and other components	Equation 1	**0.86 to 3.22**	1.1 to 4.3	Regression used over period; different methods and forcing data sets used
Otto et al. 2013	1970–2009 cf. 1860–1879	Global heat uptake estimate made for 1961–2011 from various sources; $N =$ 0.08 Wm^{-2} for base period	Equation 2	1.72	**2**	Evaluated different periods; used an ERF of 2 × CO_2 of 3.44 Wm^{-2}; forcings from Forster et al. (2013)
Bengtsson & Schwartz 2013	1970–2010	1970–2010 estimate based on OHC from Levitus et al. (2012), including deep ocean and other components	Equation 1	**1.83**	**2.0**	Regression used as in Schwartz (2012); various forcing data sets used
Lewis & Curry 2014	1995–2011 cf. 1859–1882	Final period, same source as Otto et al. (2013); $N =$ 0.15 Wm^{-2} for base period, from modeled steric sea-level rise, scaled down	Equation 2	2.25	**1.64**	Various sensitivity tests performed; AR5 forcings used (Myhre et at. 2013)
Kummer & Dessler 2014	1958–2010 cf. 1880–1900	1958–2010 total OHC reanalyses from Balmaseda et al. (2013) plus other terms; implicit zero base period assumed	Equation 2	**1.1 to 1.6**	2.3 to 3.4	Range depends on transient forcing efficacy of aerosol; AR5 forcings used

(*Continued*)

Table 1 *(Continued)*

Study	Year range	Energy imbalance (*N*)	Model	α (Wm^{-2} K^{-1})	ECS (K)	Notes
Masters 2014	1955–2011	1955–2011 different OHC data sets used, including Domingues et al. (2008)	Equation 2	**2.05**	**1.98**	Various base periods and time periods analyzed within dates; GISS forcings used (Hansen et al. 2007).
Short-timescale (decadal) analyses						
Tsushima et al. 2005	1985–1999	ERBE 60°S–60°N	Equation 1, no forcing	**0.98**	**3.8**	Overestimate of solar sensitivity by factor of 2
Forster & Gregory 2006	1985–1996	ERBE 60°S–60°N	Equation 1	**2.3**	1.6	Effects of regression explored
Lindzen & Choi 2009	1985–1999	Tropical ERBE	Equation 1, no forcing	**4.5**	0.8	Several errors in method identified
Murphy et al. 2009	1985–2005	ERBE and CERES	Equation 1	**1.25**	3.0	Explored seasonal and interannual analyses
Trenberth et al. 2010	1985–1999	60°S–60°N ERBE	Equation 1	**0.8 to 1.6**	2.3 to 4.6	Range depends on case investigated
Chung et al. 2010	1985–1999	60°S–60°N ERBE	Equation 1, no forcing	**0.11**	34	Very big shortwave feedback found
Lindzen & Choi 2011	1985–2008	Tropical ERBE and CERES from 2000	Equation 1, no forcing	**6.9**	0.5	Similar errors as in their earlier work
Tsushima & Manabe 2013	1985–2005	1985–1999, 60°S–60°N ERBE and global CERES from 2000	Equation 1, no forcing	**1.1**	3.4	Used a gain factor approach
Dessler 2013	2000–2010	Reanalyses	Sum of feedbacks	**1.15**	3.2	—
Donohoe et al. 2014	2000–2013	CERES global data	Equation 1	**1.2**	3.1	—
Trenberth et al. 2015	2000–2013	CERES global data	Equation 1, no forcing	**1.13**	3.3	$\alpha = 2.28$ Wm^{-2} K^{-1} if using tropospheric temperatures for regression
Climate model–based analyses (long timescale)						
Forster & Taylor 2006	70 years of 1% per year CO_2 increase	Model result	Equation 1	**1.42**	2.7	Average of 20 CMIP3 models
Soden & Held 2006	2000–2100	SRES A1B scenario	Sum of feedbacks	**1.28**	2.9	Average of 12 CMIP3 models
Forster et al. 2013 Andrews et al. 2012	$4 \times CO_2$ abrupt runs	Model result	Equation 1	**1.13**	**3.32**	Average of 23 CMIP5 models

(*Continued*)

Table 1 ***(Continued)***

Study	Year range	Energy imbalance (N)	Model	α (Wm^{-2} K^{-1})	ECS (K)	Notes
Dessler 2013	200 years of either preindustrial or CO_2 increase	Model result	Sum of feedbacks	**0.6 (unforced) or 1.26 (forced)**	2.9 or 6.1	Average of 13 CMIP5 models; difference from unforced and forced runs; higher α in models with good ENSO
Masters 2014	1955–2011	Model result	Equation 1	**1.18**	3.1	Assumed aerosol forcing; 32 CMIP5 models

Bold entries are taken directly from the paper; others are estimated using ECS $= F_{2\times CO_2}/\alpha$, assuming $F_{2\times CO_2} = 3.7\ Wm^{-2}$. Data sources: OHC data, Levitus et al. (2012) or earlier incarnation of same data set referenced within; ERBE, Wong et al. (2006); CERES, Loeb et al. (2009). Abbreviations: AR5, Intergovernmental Panel on Climate Change Fifth Assessment Report; CERES, Clouds and the Earth's Radiant Energy Systems; CMIP, Coupled Model Intercomparison Project; ECS, equilibrium climate sensitivity; ENSO, El Niño Southern Oscillation; ERBE, Earth Radiation Budget Experiment; ERF, effective radiative forcing; GISS, Goddard Institute for Space Studies; OHC, ocean heat content; SRES, Special Report on Emissions Scenario.

review concerns itself with understanding differences between the best estimates of the various studies. The long-timescale analyses tend to quote ECS and/or corresponding α values, whereas most of the short-timescale analyses explicitly state that their values of α are not necessarily representative of long-term climate change and therefore do not attempt to derive an ECS value. Note that emboldened numbers in **Table 1** (filled symbols in **Figure 1**) are quoted directly in the cited paper. The long- and short-timescale analyses have different pedigrees in the literature and are discussed separately in the following section.

4. EQUILIBRIUM CLIMATE SENSITIVITY FROM MULTIDECADAL DATA

The studies outlined in **Table 1** can be considered successors of the study by Gregory et al. (2002) that employ Equation 1 or 2, or a close variant, to determine ECS. These form a subset of a wider group of studies that employ some form of simple energy budget model that has climate sensitivity and/or ocean heat uptake as tunable parameters. Historically such approaches have been used to simultaneously constrain both aerosol radiative forcing and climate sensitivity (Forest et al. 2006, Knutti et al. 2002). More recent examples of this approach have been able to exploit both tighter bottom-up constraints on aerosol forcing (Boucher et al. 2013) and better estimates of ocean heat content (Rhein et al. 2013). Employing these constraints, simple model approaches give a similar range of sensitivities to those seen in **Table 1** and have best estimates of ECS less than 3.0 K. Johansson et al. (2015) found a best estimate of 2.5 K. Their ECS values are at the high end of the range. Other studies have best estimates below 2.0 K: Skeie et al. (2014) found a best estimate of 1.8 K and Lewis (2013) a best estimate of 1.6 K. Such studies are somewhat comparable to those listed in **Table 1**, but for the sake of simplicity we focus on those that use Equation 1 or 2 more directly.

We investigate the uncertainty or potential bias in the T, N, and F values employed, beginning by examining how well we can close Earth's energy budget using recent best estimates. **Figure 2** presents the current state of this closure employing data taken from AR5. This gives an illustrative best estimate of α around 1.8 Wm^{-2} K^{-1}, corresponding to an ECS around 2 K, which fits within the canonical ECS in AR5 and earlier assessments. This shows that estimates of effective

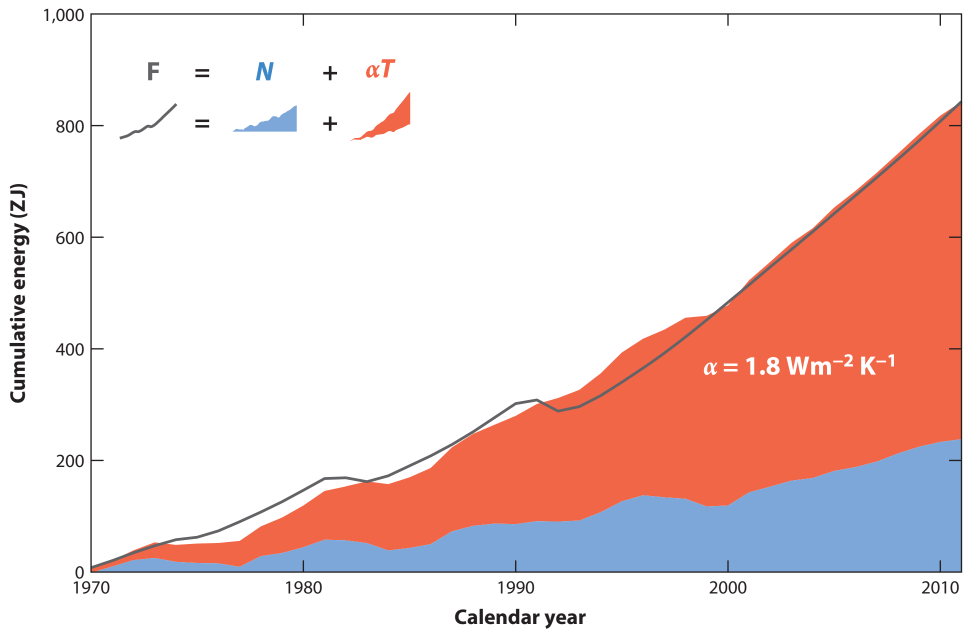


Figure 2

This graph illustrates the closure of Earth's energy budget with a rearrangement of Equation 1. Cumulative energy totals of N and F are shown integrated over 1970–2011 in units of zettajoules (ZJ; 1 ZJ = 10^{21} J). These are taken from chapters 3 (Rhein et al. 2013) and 8 (Myhre et al. 2013) of the Intergovernmental Panel on Climate Change Fifth Assessment Report (AR5) (Stocker et al. 2013), respectively. Hadley Centre–Climatic Research Unit version 4 (HadCRUT4) data (Morice et al. 2012) are used for T. The F and T baselines are taken as the 1860–1879 average. An N baseline of 0.08 Wm^{-2} is assumed (Lewis & Curry 2014, Otto et al. 2013). These baselines are subtracted from the 1970–2011 data. The energy emitted to space from the integral of αT is estimated employing an α of 1.8 Wm^{-2} K^{-1}. Other data source choices and/or different base periods lead to a range of best estimates.

sensitivity, derived from values of N, F, and T that have been measured and/or inferred from disparate observations, give a meaningful value of the ECS that is in broad agreement with other studies. Provided data sets can be considered independent, this result should be regarded as an important test that our observational capability and physical theory have passed, illustrating the fundamental robustness of climate science (see Church et al. 2013).

The cumulative approach used in **Figure 2** effectively compares N, T, and F integrated over 1970–2011 with their preindustrial counterparts. This is just one method based on Equation 1 to determine climate sensitivity. When employing such methods, uncertainty arises from a number of sources, discussed below and illustrated in **Figure 3**.

4.1. Method and Periods Analyzed

Some studies are relatively simple and compare F, N, and T over a recent period to an assumed zero point during preindustrial times (Kummer & Dessler 2014, Lin et al. 2010). However, most choose their analysis periods carefully and perform a degree of sensitivity testing. Although all methods are based on Equation 1 the exact calculation technique differs, from the use of mixed layer models (Lin et al. 2010, Schwartz 2007) to a regression approach (Bengtsson & Schwartz

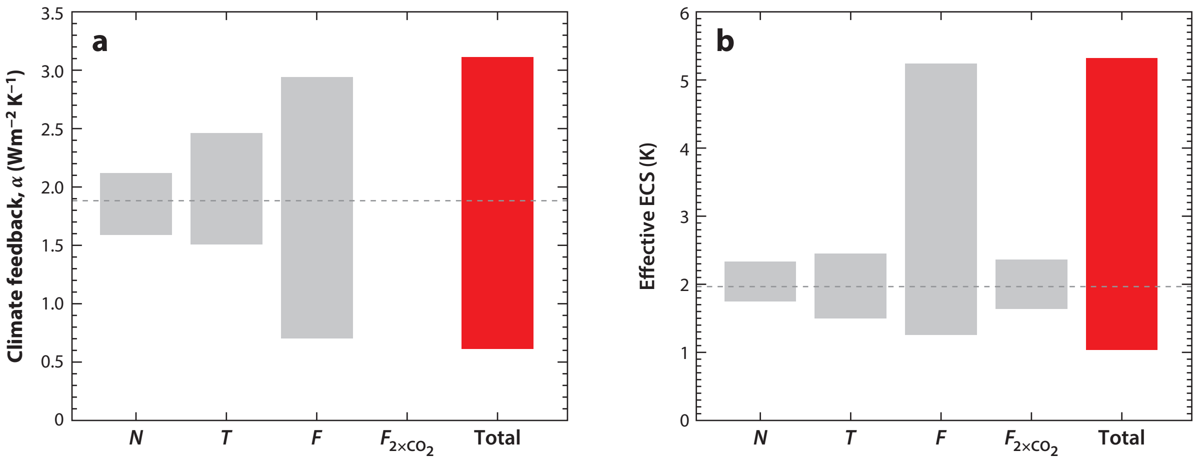


Figure 3

Uncertainty analysis for (*a*) the climate feedback parameter, α, and (*b*) effective equilibrium climate sensitivity (ECS) diagnosed from the historical record. Equation 1 and 2 are used to derive α and effective ECS, respectively, comparing present-day values to preindustrial values. Upper- and lower-bound 5% and 95% uncertainty ranges are assumed for N, T, and F and their effect on α estimated using Equation 1. The additional uncertainty in the effective radiative forcing of $F_{2\times CO_2}$ is taken into account in the effective ECS calculation (Equation 2). The uncertainty ranges are based on the Intergovernmental Panel on Climate Change Fifth Assessment Report (AR5) (Stocker et al. 2013), and errors are assumed to be Gaussian and are added in quadrature to estimate the total uncertainty. The best estimates and 5–95% ranges are as follows: N, 0.6 Wm^{-2} (0.4 to 0.85 Wm^{-2}); T, 0.85 K (0.65 to 1.06 K); F, 2.2 Wm^{-2} (1.2 to 3.1 Wm^{-2}); $F_{2\times CO_2}$, 3.7 Wm^{-2} (3.1 to 4.4 Wm^{-2}). N is estimated from the 0.72 Wm^{-2} 2000–2011 energy imbalance calculated from data in chapter 3 of AR5 (Rhein et al. 2013), applying a preindustrial offset of 0.1 Wm^{-2} for the best and 5% range estimates and zero for the 95% range estimate. T and F are anomalies computed from the 1880–2012 linear trends and their uncertainty. T trends are quoted from chapter 2 of AR5 (Hartmann et al. 2013). F trends and $F_{2\times CO_2}$ are taken from chapter 8 of AR5 (Myhre et al. 2013). Our simple analysis overestimates the contribution of $F_{2\times CO_2}$ to the overall uncertainty as it assumes $F_{2\times CO_2}$ would be independent of F, when in reality there is a degree of compensation (Lewis & Curry 2014).

2013, Schwartz 2012) to time period comparisons (Gregory et al. 2002, Roe & Armour 2011, Lewis & Curry 2014, Masters 2014, Otto et al. 2013) to an integral approach as seen in **Figure 2** (Kummer & Dessler 2014). The integral approach has been adopted to reduce the effects of noise/variability. Other studies achieve a similar outcome by averaging N, T, and F over sufficiently long present-day and base periods (Lewis & Curry 2014, Otto et al. 2013). From conservation of energy, a variability-driven change in one variable should drive consistent changes in the others. Nevertheless, considering variability is important, as variability-driven changes in the geographical pattern of surface temperature (for example) could cause an instantaneous change in feedback (α values), driving it away from its long-term mean (Colman & Power 2010). Such variability might even affect effective radiative forcing (ERF) (Lewis & Curry 2014). Likewise, it is desirable to avoid volcanic periods, as feedbacks from volcanic forcing are likely dissimilar from those due to greenhouse gases (Forster & Gregory 2006, Lewis & Curry 2014, Merlis et al. 2014). Lewis & Curry (2014) and Otto et al. (2013) showed how poor choices over base periods can lead to biases in ECS of a few tenths of a degree. Masters (2014) took a different approach and evaluated many different base and "current" periods between 1955 and 2011, and thereby avoided the need to make an estimate of preindustrial baselines. The approaches taken by Schwartz (2012) and Bengtsson & Schwartz (2013) avoided the need for a baseline by using regression over the past 50 years, a similar technique to that employed in the short-timescale analyses. These techniques have their own issues (see Section 5). Schwartz (2007) employed the most novel technique, avoiding the need for a forcing term by using N and T data sets to determine both the heat capacity and time constant of the Earth system through an analysis of autocorrelation. This method derives a low value of ECS, but it is doubtful whether a model of Earth with a fixed heat capacity and time constant is appropriate (Flato et al. 2013).

4.2. Earth's Energy Imbalance Data

The primary data come from measurements of ocean temperature changes in the top 700 m of the ocean. Satellite-based estimates, when used, are typically matched to ocean heat-content changes (Loeb et al. 2012). For a more globally representative energy imbalance, most estimates in **Table 1** combine upper ocean data with deep ocean heat-content data and other assumptions regarding heat-content changes from nonocean reservoirs. Levitus et al. (2012) provide ocean heat-content data between 700 and 2,000 m depth; data reported by Balmaseda et al. (2013) and in AR5 also take heating within the abyssal ocean into account. The AR5 data add components of heating due to ice melt and changes to the land and atmosphere; see box 3.1 in chapter 3 of AR5 (Rhein et al. 2013). Data from AR5 and Domingues et al. (2008) indicate a present-day imbalance of around 0.7 Wm^{-2}. Box 3.1 in chapter 3 of AR5 (Rhein et al. 2013) gives a 24% uncertainty range for the 1993–2010 heat uptake, which would translate to a range of N values that is very similar to that used in the studies described in **Table 1**. For completeness many studies also account for a small energy imbalance of around 0.1 Wm^{-2} at the end of the nineteenth century that has been identified within climate models (Gregory et al. 2002, Lewis & Curry 2014, Otto et al. 2013). This acts to slightly reduce N, leading to a slightly smaller estimate of ECS. Differences in the assumed N account for most of the spread in sensitivity estimates between the studies listed in **Table 1** (see also **Figure 3** and Lewis & Curry 2014). Examining estimates on N for the extreme ECS estimates identified in **Table 1**, we find that Lewis & Curry (2014) assumed a present-day (~2000–2009) imbalance of 0.51 Wm^{-2}, and a preindustrial offset was applied to give a relatively small estimate of the change in N. In contrast, Kummer & Dessler (2014) and Lin et al. (2010) assumed an imbalance greater than 0.7 Wm^{-2} and did not apply a preindustrial offset to estimate N. We evaluated energy imbalance trends from the Coupled Model Intercomparison Project phase 5 (CMIP5) models for this paper and found them to be similar to those employed in the long-timescale analyses in **Table 1**. Uncertainty in both the present and baseline energy imbalance contributes an uncertainty of roughly ± 0.3 K to the overall ECS estimate diagnosed from the long-term 1880–2012 record (**Figure 3**).[1]

4.3. Choice of Temperature Data

The data sets of global temperature have slight differences in their trends caused by a number of factors (Hartmann et al. 2013, Karl et al. 2015). Two uncertainties are particularly important in the context of estimating ECS. There is difficulty in defining (*a*) a preindustrial baseline and (*b*) the degree to which the 1998–2014 apparent hiatus in surface warming is a manifestation of data errors in the observations. The Arctic has warmed considerably since 2000, but some data sets, such as Hadley Centre–Climatic Research Unit version 4 (HadCRUT4), may not have adequately captured its contribution to global trends (Cowtan & Way 2014). Comparing the National Oceanic and Atmospheric Administration trends to the standard HadCRUT4 data without the correction

[1]The change in N was quantified from the climate models, comparing their 2000–2012 energy imbalance with their 1860–1880 energy imbalance, using integrations of the RCP6.0 scenario from the 25 model. Across the models the median N value was 0.67 Wm^{-2}, with a 0.33 to 1.04 Wm^{-2} 5% to 95% uncertainty range.

applied would slightly increase ECS estimates (Johansson et al. 2015). Only one temperature data set extends prior to 1880 due to the poor spatial coverage in the early record, which makes it difficult to estimate trends from a preindustrial baseline. Accounting for both these sources of uncertainty, AR5 placed an overall uncertainty in the 1880–2012 linear trend of 0.85 K with a 5% to 95% uncertainty range of 0.65 to 1.06 K; this leads to an uncertainty of roughly ± 0.5 K in ECS diagnosed from long-term trends (**Figure 3**).

4.4. Choice of Forcing Data

Uncertainty in the radiative forcing, and particularly that due to aerosol and aerosol cloud interactions, is widely taken to be the single biggest contributor to uncertainty in ECS diagnosed from the historical record (Bindoff et al. 2013, Stevens 2015). This uncertainty does not manifest itself in the range of best estimates shown in **Table 1**, as most recent studies employ similar best estimates of radiative forcing changes, taken either from AR5 (Myhre et al. 2013), CMIP5 (Forster et al. 2013), or Goddard Institute for Space Studies (GISS) (Hansen et al. 2007) data analysis. Aside from aerosols, there are four other important sources of uncertainty to consider:

1. **The level of volcanic forcing.** There were several large explosive volcanic eruptions in the early record, such as the eruption of Krakatau in 1883. These were of uncertain magnitude and affect the overall forcing trend. The α associated with volcanic response may also be quite different than that due to CO_2 (Forster & Gregory 2006). Lewis & Curry (2014) managed this uncertainty by choosing to compare present and past periods of similar volcanic forcing (see Section 4.1) and/or by assuming a reduced efficacy for volcanic forcing (see item 3, below).
2. **The effect of rapid adjustments.** Equation 1 splits the energy budget into one part associated with the forcing and another part due to global surface temperature change; therefore, any change in imbalance caused by forcing that is unrelated to global temperature change needs to be accounted for in F. This means that F not only accounts for the traditional stratospheric adjusted radiative forcing as defined in AR4 (Forster et al. 2007), but also needs to include any rapid adjustments to this forcing that would affect the top-of-atmosphere energy imbalance (Boucher et al. 2013, Sherwood et al. 2015). Rapid adjustments could come from atmospheric stability changes affecting clouds and/or changes in patterns of land-surface heating. Such adjustments have been calculated in models for an increase in carbon dioxide and some aerosol changes but not for the other forcing terms (Sherwood et al. 2014). Further, we have as yet not been able to find real-world measurements of rapid adjustment to check the model response. Generally, rapid adjustments could add both uncertainty and systematic bias to the evaluation of F.
3. **The role of efficacy in evaluating the forcing response.** An uncertainty closely related to rapid adjustments is the role of efficacy, which tells one how effectively a given watts per square meter forcing by one mechanism triggers a global mean temperature response, compared with a watts per square meter forcing from a CO_2 increase. Efficacies are used to account for how α might vary across different forcing mechanisms. By including rapid adjustments efficacies should be closer to 1, making the assumption of using a single α more applicable (Forster et al. 2013, Sherwood et al. 2015). However, this is not true of very regionally confined forcings, such as black carbon on snow (Bond et al. 2013). The efficacy of volcanic forcing could be less than 1, and this could lead to a possible overestimate of ECS if analysis periods are not chosen carefully (Lewis & Curry 2014). In contrast, Shindell (2014) and Shindell et al. (2015) have suggested that aerosol forcing may have a larger transient response or short-term efficacy, due to cooling Northern Hemisphere land surfaces at a faster

rate than a more globally distributed forcing such as CO_2 might. Ignoring such an aerosol efficacy would underestimate the ECS diagnosed from the historical record (Kummer & Dessler 2014). This theory is expanded in Section 4.5.

4. **The assumed radiative forcing for a doubling of carbon dioxide.** In calculating ECS from α, a forcing for a doubling of CO_2 needs to be applied. $F_{2\times CO_2}$ is typically taken to be 3.7 Wm^{-2}, and its uncertainty is often ignored in estimates of ECS (Lewis & Curry 2014). Otto et al. (2013) used a lower forcing value of 3.44 Wm^{-2} to account for rapid adjustments based on model results of the ERF. More generally, Myhre et al. (2013) took the uncertainty in ERF for $F_{2\times CO_2}$ to be $\pm 20\%$. This uncertainty in the numerator of Equation 2 is somewhat compensated by CO_2 uncertainty contributing to the uncertainty in F, appearing in the denominator of Equation 2.

The forcing uncertainty range shown in **Figure 3** assumes that uncertainties from sources 1, 2, and 3 are implicitly already included within the broad AR5 forcing estimate. The role of $F_{2\times CO_2}$ uncertainty is assessed separately for ECS. Historical forcing uncertainty leads to a 50% change in α and to a broad range of possible ECS values. $F_{2\times CO_2}$ contributes a further uncertainty, but only to ECS. Unknown rapid adjustments or efficacy effects could lead to important potential sources of bias. Note that quantifying all four of the above uncertainty sources relies on climate models, emphasizing the fallacy that energy budget estimates are independent of models.

4.5. Discussion

Figure 3 uses data solely from the long-term changes quantified in AR5 to illustrate the role of various uncertainties in α and ECS. The illustrative best estimates of α and ECS are 1.88 Wm^{-2} K^{-1} and 1.97 K, respectively. The simple error analysis employs the IPCC uncertainty ranges and assumes errors are Gaussian and add in quadrature. Overall, 5% to 95% error ranges for ECS are roughly between 1.0 and 5.0 K, slightly broader than the AR5 canonical range. Recent range estimates tend to have relatively low 95% upper bounds for ECS: Otto et al. (2013) estimated a 3.9 K upper bound, Lewis & Curry (2014) estimated 4.1 K, and Kummer & Dessler (2014) estimated 4.1 K (excluding efficacy effects). By contrast, Masters (2014) determined an upper bound of 5.1 K. The analysis in **Figure 3** is cruder than that employed in these cited studies, but note that the high ECS uncertainty is very sensitive to the best estimate. For example, if we applied the same percentage uncertainty ranges to an ECS best estimate of 1.7 K, we would restrict the 95% upper bound to 3.5 K. The large variation in the 95% range estimate with the best estimate is due to the reciprocal relationship between α and ECS. This highlights the fact that the 95% upper bound for ECS from energy budget studies is not that robust.

The range of ECS presented in **Figure 3** and the best estimates from the multidecadal analyses shown in **Figure 1** and **Table 1** are consistent, but they are at the low end of the AR5 assessed range. Their ~2 K best estimates may well prove correct, especially as research is converging on a better constrained and somewhat more modest value for the globally averaged aerosol forcing (Boucher et al. 2013). Nevertheless, the best estimates and the lower bound on ECS remain considerably smaller when compared with recent estimates that employ other methods of diagnosing ECS (e.g., Sherwood et al. 2014). This discrepancy has prompted researchers to return to the question of whether it is a good approximation to assume that α over the historical period is the same as that associated with a doubling of carbon dioxide, referred to here as the effective sensitivity question.

It is clear that for large temperature changes the simple linear relationship between forcing and temperature breaks down (Caballero & Huber 2013). Nonlinearity also becomes more pronounced at high ECS (Bloch-Johnson et al. 2015). Nonlinear behavior may also appear for ~1 K global temperature changes as some models exhibit a time dependence in their effective sensitivity (Andrews et al. 2014, Armour et al. 2013, Collins et al. 2013, Winton et al. 2013). Masters (2014) had to make assumptions about historical forcings but suggested that the α derived from the IPCC Special Report on Emissions Scenario (SRES) A1B in CMIP3 models was 1.09 ± 0.2 times larger than the α for a doubling of carbon dioxide, implying that the true ECS is on average slightly higher than its effective sensitivity, but this result is not that conclusive. Andrews et al. (2014) found 23 of 27 CMIP5 models exhibit increasing effective sensitivity (smaller α) through time. An analysis of the 1% per year CO_2 integrations from 18 models suggested the true ECS could be 20–40% larger than the effective ECS (K. Armour, personal communication). The latest generation of models also show increased water vapor feedback and ECS at higher base temperatures (Meraner et al. 2013). Generally, models exhibit regional variation in feedbacks (Crook et al. 2011), and as the spatial distribution of surface temperature evolves, it affects the time evolution of α (Armour et al. 2013). Changing ocean heat uptake has also been implicated (Block & Mauritsen 2013, Geoffroy et al. 2013, Koll & Abbot 2013, Rose et al. 2014, Winton et al. 2013). Rose et al. (2014) suggested ocean heat uptake changes drive differences in shortwave cloud feedback, affecting sensitivity. Taken together, these studies suggest that ECS diagnosed from the twentieth-century record may significantly underestimate the ECS for a doubling of carbon dioxide. However, even complex atmosphere-ocean coupled climate models can be remarkably linear in their response (Andrews et al. 2012), so the degree of underestimation remains uncertain.

Shindell (2014) suggested that a higher transient efficacy of aerosol forcing could cause an underestimate of ECS in the historical record. Depending on the efficacy assumed for the aerosol forcing, the underestimate of ECS could be as much as 1.1 K (Kummer & Dessler 2014). The Shindell (2014) and Shindell et al. (2015) studies could only draw speculative conclusions due to a lack of forcing information within the historical CMIP5 model integrations analyzed. Marvel et al. (2015) explored this properly by computing efficacies and ERFs for individual forcing terms. They showed that within the GISS model, accounting for efficacies in the historical response could raise the implied best estimate of ECS from the historical period from around 2.0 K to around 3.0 K. However, my preliminary investigations (not shown) could not find any evidence of this effect in dedicated aerosol and carbon dioxide perturbation experiments within two climate models [Community Climate System Model version 4 (CCSM4) and Hadley Centre Global Environment Model version 2 (HadGEM2)]. Nevertheless, it could still be important, as the efficacies of different ERF mechanisms have not been properly assessed across models. The groundbreaking efficacy study by Hansen et al. (2005) needs updating: Efficacies need evaluating across other models and for transient change.

In summary, long-term energy budget analysis suggests a best estimate of effective ECS around 2.0 K, lower than the ECS estimates suggested by other methods. Energy budget analyses can usefully constrain α to between 0.6 and 3.1 $Wm^{-2}\ K^{-1}$ (**Figure 3**). A viable likely uncertainty range for effective ECS is 1–5 K, consistent with the IPCC range. ERF uncertainty dominates the uncertainty in both α and ECS derived from such methods. The range for ECS is broader than some of the published ranges from studies listed in **Table 1**, which have claimed relatively tight constraints on ECS using such approaches (not shown). This is principally because the high-end estimates of ECS range are not that robust due to the reciprocal relationship between ECS and α. The effective ECS diagnosed from these methods may be systematically different than that associated with long-term changes in CO_2. There is emerging evidence that it could be biased significantly toward the low end due to the way feedbacks are expected to evolve through time. This could be due to spatial evolution of feedbacks related to surface temperature pattern evolution. There is also speculation that transient aerosol cooling is underestimated in such approaches, and this could also contribute to a biased-low best estimate of ECS from historical energy budget analysis. Allowing for the translation of effective ECS to actual ECS would increase the best estimate and the range estimates shown in **Figure 3**.

5. EQUILIBRIUM CLIMATE SENSITIVITY FROM INTERANNUAL DATA

Approaches that employ shorter-term interannual data sets shown in **Table 1** have a different pedigree than the long-term energy budget analyses. They have evolved from early satellite investigations into the greenhouse effect, cloud forcing, and feedbacks (e.g., Ramanathan et al. 1989). They invariably concentrate on diagnosing α rather than estimating ECS, and all acknowledge that an α diagnosed employing a short-term record may not be representative of its longer-term value. The field is more fragmented than the longer-timescale-based estimates: Findings often sit as standalone estimates in relative isolation, as there are fewer comparisons between different published estimates. Erroneous estimates of α have been published using these approaches, and there is no agreed methodology of data analysis. Consequently, the best estimates quoted in **Table 1** cover a very wide range of α values. As in Section 4, differences arise due to the choice of data and method used. These are discussed below and illustrated in **Figure 4**.

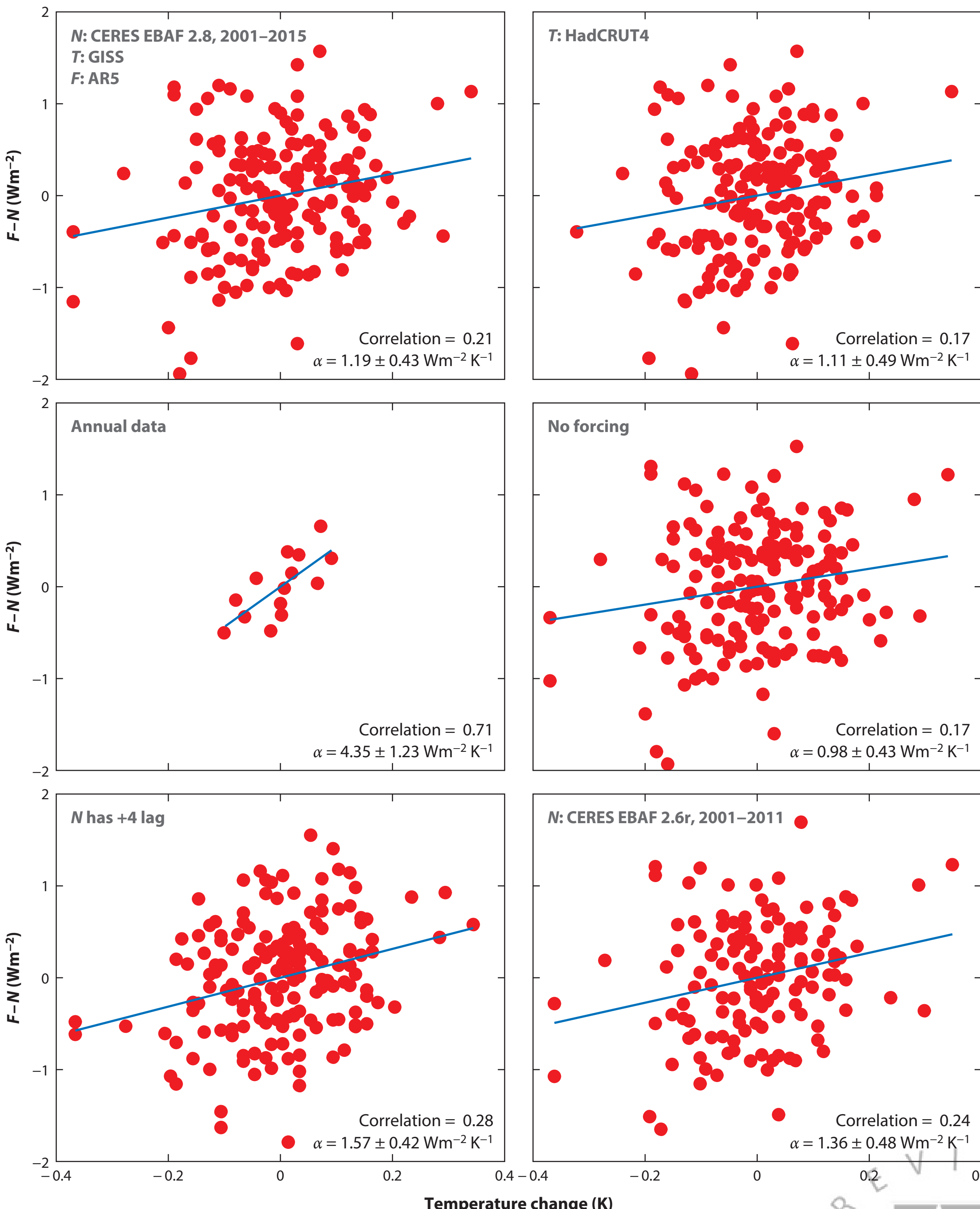
5.1. Data Choices

Unlike in the case of multidecadal change, forcing plays a much smaller role on interannual time periods unless there is significant volcanic activity. Instead time series are dominated by variability. There is also less coherence between surface temperature data sets. Surface and free tropospheric temperatures have both previously been used to derive α values (Trenberth et al. 2015).

A major source of uncertainty in α comes from uncertainty in estimates of Earth's energy imbalance and its changes through time. Interannual changes in ocean heat content are not to be trusted (Church et al. 2013), so studies depend on the net radiative flux estimates from satellites. Two related data sets exist. The Earth Radiation Budget Experiment (ERBE) instruments flew on three satellites launched in the mid-1980s, and data were collected by the nonscanner instruments between 60°N and 60°S, providing 15 years of data from 1985 to 1999 (Wong et al. 2006). These data have gone through several version controls, especially to account for the drift in shortwave radiation measurements, which greatly affected decadal variability. For example, Forster & Gregory (2006) and Murphy et al. (2009) used different versions of ERBE data when diagnosing α (**Table 1**). ERBE was followed by Clouds and the Earth's Radiant Energy Systems (CERES) instruments flying on a number of satellites beginning in 2000. These data have again undergone a series of corrections. The latest Energy Balanced and Filled data set, edition 2.8 (CERES EBAF 2.8), uses infilling to obtain global coverage and ocean heat uptake analyses to calibrate net flux

Figure 4

An illustration of how data and methodological choices affect the diagnosis of the climate feedback parameter, α. The reference estimate is shown in the top left panel using monthly 2001–2013 Clouds and the Earth's Radiant Energy Systems Energy Balanced and Filled edition 2.8 (CERES EBAF 2.8) data (Loeb et al. 2009), Goddard Institute for Space Studies (GISS) temperature data (Hansen et al. 2007), and Intergovernmental Panel on Climate Change Fifth Assessment Report (AR5) forcing data (Stocker et al. 2013). Other panels change either one element of the data (*right panels*), or one element of the method (*left panels*). In this instance the maximum correlation was found when N lagged the T anomaly by 4 months (*bottom right panel*). The blue line in each panel shows the least squares linear regression straight line fit to the data. The gradient of this fit gives the estimate of α and its uncertainty indicated in the bottom right of the panel.


N: CERES EBAF 2.8, 2001–2015
T: GISS
F: AR5
Correlation = 0.21
α = 1.19 ± 0.43 Wm⁻² K⁻¹
T: HadCRUT4
Correlation = 0.17
α = 1.11 ± 0.49 Wm⁻² K⁻¹
Annual data
Correlation = 0.71
α = 4.35 ± 1.23 Wm⁻² K⁻¹
No forcing
Correlation = 0.17
α = 0.98 ± 0.43 Wm⁻² K⁻¹
N has +4 lag
Correlation = 0.28
α = 1.57 ± 0.42 Wm⁻² K⁻¹
N: CERES EBAF 2.6r, 2001–2011
Correlation = 0.24
α = 1.36 ± 0.48 Wm⁻² K⁻¹
F–N (Wm⁻²)
Temperature change (K)
2
1
0
−1
−2
−0.4
−0.2
0
0.2
0.4

estimates (Loeb et al. 2009). Because of data revisions and improvements, and due to the increase in the record length, estimates from studies that employ the latest data sets may be more robust than some earlier estimates.

5.2. Methods

An important consideration is that data should be as near global as possible. Studies based on tropical analyses (Lindzen & Choi 2009, 2011; Spencer & Braswell 2008) are confounded by meridional heat transport and do not provide a representative estimate of α (Chung et al. 2010, Murphy 2010).

Estimates of α are typically derived from a regression analysis of N versus T, or by taking differences in N and T across two periods. Choosing specific periods to compare (Lindzen & Choi 2009, 2011) is prone to bias and can lead to unrepresentative estimates of α (Trenberth et al. 2010). Another error in several studies is that they neglect the role of forcing entirely (see **Table 1**); this leads to biases in diagnosing α (Chung et al. 2010, Murphy & Forster 2010). It is especially important to include ERF trends when considering volcanic periods (Forster & Gregory 2006). However, volcanic forcing has changed little since 2000. When analyzing this period, excluding the forcing term likely leads to only minimal bias in α, less than 0.1 $\mathrm{Wm^{-2}\ K^{-1}}$, and making this approximation can be seen as a legitimate choice (Trenberth et al. 2015). A number of other specific analysis errors were also identified in the Lindzen & Choi (2009, 2011) publications (Dessler 2011, Trenberth et al. 2010).

Generally studies with obvious errors in approach can be ignored. However, other methodological choices also affect results and possibly create biases. These choices principally concern themselves with treatment of the seasonal cycle and the method used for regression. Diagnosing α from regression using Equation 1 ignores the possibility that random changes in flux (γ) might also change T. This leads to a systematic underestimation of α (Spencer & Braswell 2008), although any underestimation is likely small ($\sim$0.05 $\mathrm{Wm^{-2}\ K^{-1}}$) for realistic cases (Murphy & Forster 2010). Flux changes are typically regressed against surface temperatures, as to first order it is likely that surface temperature changes drive top-of-atmosphere fluxes on these timescales. However, as energy imbalance changes also drive temperature changes, the argument for this regression choice is not clear cut. Using total least squares regression, for example, would lead to a significant increase in estimates of α (Forster & Gregory 2006). Likewise, including or excluding the seasonal cycle leads to different values of α (Forster & Gregory 2006, Murphy et al. 2009). Lagging the N data before regressing it against T might improve estimates, but this approach has not been properly explored.

The two recent estimates of α from short-timescale analysis both use monthly CERES data with the annual cycle removed, assuming zero lag. Donohoe et al. (2014) include the F term, but Trenberth et al. (2015) do not. The two groups also employ slightly different surface temperature data. Nevertheless, both studies find similar α values, around 1.2 $\mathrm{Wm^{-2}\ K^{-1}}$ (**Table 1**). These are also similar to α values derived from annual data using a slightly different gain factor method (Tsushima & Manabe 2013) and to those derived from using regression to evaluate feedbacks in reanalysis data (Dessler 2013).

5.3. Discussion

Figure 4 illustrates how some of the data and methodological choices affect α. Choosing different data for T, N, and F affects estimates of feedback. The diagnosed feedback is relatively robust to choice of T and F. Lagging N by a few months appears to slightly improve the correlation, but this is not a strong dependence. It is not well understood why annual regressions give very different α values. This could be an artifact of the short time series, a result of different physical feedbacks operating on intraannual and interannual time series, or an issue with the CERES time series. Unlike the long-term energy budget changes, forcing uncertainty does not dominate the derivation of α from short-term energy budget changes. The methodological choice accounts for most of the spread between published best estimates (**Table 1** and **Figure 4**). As there is only one quality-controlled measure for interannual variations in N, it is difficult to accurately gauge its contribution to uncertainty. For comparison, an earlier incarnation of the CERES data (edition 2.6r) that spans 2001–2011 is shown. This would indicate a relatively small uncertainty arising from the N term. However, the true contribution from uncertainty in N is likely considerably larger. Having other groups analyze the satellite record to make an independent estimate of N would be extremely useful.

Model feedback analyses and regression analyses over longer periods of model data for various models and scenarios are also presented in **Table 1**. These studies find long-term α values between 0.6 and 1.4 $\mathrm{Wm^{-2}\,K^{-1}}$. These model results appear to agree with the recent estimates of α derived from the observations. However, as the relationship between short-timescale and long-timescale feedbacks has not yet been thoroughly tested in models, the apparent agreement may be fortuitous (see Section 6).

It is unclear how to relate the short-term values of α to their longer-term or equilibrium counterparts. Dessler (2013) suggested that 10 years of data may be sufficient. Gordon et al. (2013) examined CMIP5 models and found that 10 years of data is really the minimum needed for a regression-based estimate of a given model's water vapor feedback to agree with its long-term values; for some models several decades of data were required. In the CERES data, as the annual data is better correlated than the monthly data with surface temperature and the types of averaging give such different sensitivities, it is hard to be convinced that recent published estimates of α are representative of the long-term response even though they appear to quantitatively agree with model analyses. In ECS terms, annual data would suggest an ECS greater than 1 K and monthly data an ECS greater than 3 K. Without understanding why these are so different, it becomes impossible to gauge which to trust more.

6. THE WAY FORWARD

There are multi-million-dollar work programs under way that continually improve our observational capability. Continuous Argo float measurements are expanding to include measurements at ocean depths below 2,000 m to properly understand ocean heat uptake. Refinements to CERES satellite products are beginning to constrain the net top-of-atmosphere radiation imbalance. Likewise, research groups around the world are improving estimates of surface temperature trends and working hard on constraining aerosol forcing. These efforts should help improve the data input into energy budget analyses of ECS. The most pressing need is to translate estimates of effective ECS derived from such energy budget analyses into the actual ECS. We need to remember that models are needed to make this translation, so no estimate employing observations can be said to be independent of a climate model.

Masters (2014) made the most sophisticated attempt to date to evaluate the use of Equation 1 in models. Masters compared α values derived from Equation 1 over the past 50 years with the α value derived by comparing the preindustrial period with 2100 under IPCC Representative Concentration Pathway 4.5 (RCP4.5) (following Soden & Held 2006). As forcing data for either period were not available from the models, he assumed models had identical forcings taken from standard data sets. With these assumed forcings, a reasonably good agreement between the

THE RADIATIVE FORCING MODEL INTERCOMPARISON PROJECT

The Radiative Forcing Model Intercomparison Project (RFMIP) is part of the World Climate Research Programme's Coupled Model Intercomparison Project phase 6 (CMIP6). It is being led by Robert Pincus, Bjorn Stevens, and myself. Modeling groups are encouraged to sign up and commit to providing effective radiative forcing (ERF) estimates from their models. Details can be found at **http://www.wcrp-climate.org/modelling-wgcm-mip-catalogue/modelling-wgcm-mips/418-wgcm-rfmip**.

short- and long-term α values in found. However, the derived α is slightly larger for the longer period (1.44 Wm^{-2} K^{-1}) than for the 50-year period (1.18 Wm^{-2} K^{-1}).

Masters' (2014) study was not a perfect model test, as it had to make gross assumptions about the forcing. Because of the move to interactive aerosol schemes, CMIP3 and CMIP5 models did not typically perform double radiation calls, and their forcings had to be backed out from inverting Equation 1 assuming constant α (Forster et al. 2013, Forster & Taylor 2006). Neglecting to output a forcing diagnostic severely limits the use of models for understanding important aspects of climate response. Perfect model tests would allow many outstanding questions over approaches for diagnosing ECS over the instrumental record to be swiftly laid to rest. In contrast to the effort that is being put into improving observations, the request is simple: All we need to do is to diagnose historical ERF changes from the models. As rapid adjustments need to be included in these estimates, the most straightforward way to diagnose ERF is from fixed sea-surface temperature experiments, where forcings are changed (Sherwood et al. 2015). There will be a concerted effort for CMIP6 to diagnose forcings within the models: the Radiative Forcing Model Intercomparison Project (RFMIP) (see sidebar). Having this forcing information would allow us to thoroughly test methods of diagnosing ECS from the historical record and lay many skeletons to rest.

A second important need is to understand rapid adjustments and ERF in more detail, including any forcing-related efficacy. Rapid adjustments to clouds may well be important for nonaerosol forcings, and these will influence how forcings affect Earth's energy budget. The ERF for explosive volcanic eruptions, for example, could be quite different, and possibly much smaller, than its stratospherically adjusted forcing (Gregory et al. 2016). The ERF for other forcings, including CO_2, has not been properly evaluated to date. RFMIP will estimate some of these adjustments with CMIP6 models, but we need high-resolution model tests and cleverly designed observational analyses and/or campaigns to test the models' representation of rapid adjustments and forcing in general.

The ECS diagnosed from the historical record may point to a low best estimate, but it could also simply mean that our forcing estimates are wrong or that sensitivity increases with time. A lack of knowledge about ERF and our remiss in diagnosing it within climate models are hampering progress. Improved ERF knowledge will enable much tighter constraints on ECS, and TCR will help create a robust basis for climate policy decision-making. The way forward is clear, and it would be a dereliction of duty if future model integrations did not diagnose forcings and were solely used to determine overall climate response.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

I thank Kyle Armour, Aaron Donohoe, and Nic Lewis for helpful discussions and comments on an early draft of the manuscript. Funding support was provided by a Royal Society Wolfson Merit Award and Engineering and Physical Sciences Research Council Grant EP/I014721/1.

LITERATURE CITED

Aldrin M, Holden M, Guttorp P, Skeie RB, Myhre G, Berntsen TK. 2012. Bayesian estimation of climate sensitivity based on a simple climate model fitted to observations of hemispheric temperatures and global ocean heat content. *Environmetrics* 23:253–71

Andrews T, Forster PM, Boucher O, Bellouin N, Jones A. 2010. Precipitation, radiative forcing and global temperature change. *Geophys. Res. Lett.* 37:L14701

Andrews T, Gregory JM, Webb MJ. 2014. The dependence of radiative forcing and feedback on evolving patterns of surface temperature change in climate models. *J. Clim.* 28:1630–48

Andrews T, Gregory JM, Webb MJ, Taylor KE. 2012. Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models. *Geophys. Res. Lett.* 39:L09712

Andronova NG, Schlesinger ME. 2001. Objective estimation of the probability density function for climate sensitivity. *J. Geophys. Res.* 106(D19):22605–11

Armour KC, Bitz CM, Roe GH. 2013. Time-varying climate sensitivity from regional feedbacks. *J. Clim.* 26:4518–34

Arrhenius S. 1896. On the influence of carbonic acid in the air upon temperature of the ground. *Philos. Mag. J. Sci.* 41:237–76

Balmaseda MA, Trenberth KE, Källén E. 2013. Distinctive climate signals in reanalysis of global ocean heat content. *Geophys. Res. Lett.* 40:1754–59

Bengtsson L, Schwartz SE. 2013. Determination of a lower bound on Earth's climate sensitivity. *Tellus B* 65:21533

Bindoff NL, Stott PA, AchutaRao KM, Allen MR, Gillett N, et al. 2013. Detection and attribution of change: from global to regional. See Stocker et al. 2013, pp. 867–952

Bloch-Johnson J, Pierrehumbert RT, Abbot DS. 2015. Feedback temperature dependence determines the risk of high warming. *Geophys. Res. Lett.* 42:4973–80

Block K, Mauritsen T. 2013. Forcing and feedback in the MPI-ESM-LR coupled model under abruptly quadrupled CO_2. *J. Adv. Model. Earth Syst.* 5:676–91

Bond TC, Doherty SJ, Fahey DW, Forster PM, Berntsen T, et al. 2013. Bounding the role of black carbon in the climate system: a scientific assessment. *J. Geophys. Res. Atmos.* 118:5380–552

Boucher O, Randall D, Artaxo P, Bretherton C, Feingold G, et al. 2013. Clouds and aerosols. See Stocker et al. 2013, pp. 571–658

Caballero R, Huber M. 2013. State-dependent climate sensitivity in past warm climates and its implications for future climate projections. *PNAS* 110:14162–67

Charney JG, Arakawa A, Baker DJ, Bolin B, Dickinson RE, et al. 1979. *Carbon Dioxide and Climate: A Scientific Assessment*. Washington, DC: Natl. Acad. Sci.

Chung ES, Soden BJ, Sohn BJ. 2010. Revisiting the determination of climate sensitivity from relationships between surface temperature and radiative fluxes. *Geophys. Res. Lett.* 37:L10703

Church JA, Clark PU, Cazenave A, Gregory JM, Jevrejeva S, et al. 2013. Sea level change. See Stocker et al. 2013, pp. 1137–216

Collins M, Knutti R, Arblaster J, Dufresne JL, Fichefet T, et al. 2013. Long-term climate change: projections, commitments and irreversibility. See Stocker et al. 2013, pp. 1029–136

Colman RA, Power SB. 2010. Atmospheric radiative feedbacks associated with transient climate change and climate variability. *Clim. Dyn.* 34:919–33

Cowtan K, Way RG. 2014. Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. *Q. J. R. Meteorol. Soc.* 140:1935–44

Crook JA, Forster PM, Stuber N. 2011. Spatial patterns of modeled climate feedback and contributions to temperature response and polar amplification. *J. Clim.* 24:3575–92

Dessler AE. 2011. Cloud variations and the Earth's energy budget. *Geophys. Res. Lett.* 38:L19701
Dessler AE. 2013. Observations of climate feedbacks over 2000–10 and comparisons to climate models. *J. Clim.* 26:333–42
Domingues CM, Church JA, White NJ, Gleckler PJ, Wijffels SE, et al. 2008. Improved estimates of upper-ocean warming and multi-decadal sea-level rise. *Nature* 453:1090–93
Donohoe A, Armour KC, Pendergrass AG, Battisti DS. 2014. Shortwave and longwave radiative contributions to global warming under increasing CO_2. *PNAS* 111:16700–5
Flato G, Marotzke J, Abiodun B, Braconnot P, Chou SC, et al. 2013. Evaluation of climate models. See Stocker et al. 2013, pp. 741–866
Forest CE, Stone PH, Sokolov AP. 2006. Estimated PDFs of climate system properties including natural and anthropogenic forcings. *Geophys. Res. Lett.* 33:L01705
Forster PM, Andrews T, Good P, Gregory JM, Jackson LS, Zelinka M. 2013. Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models. *J. Geophys. Res. Atmos.* 118:1139–50
Forster PM, Gregory JM. 2006. The climate sensitivity and its components diagnosed from Earth Radiation Budget data. *J. Clim.* 19:39–52
Forster PM, Ramaswamy V, Artaxo P, Berntsen T, Betts R, et al. 2007. Changes in atmospheric constituents and in radiative forcing. See Solomon et al. 2007, pp. 129–234
Forster PM, Taylor KE. 2006. Climate forcings and climate sensitivities diagnosed from coupled climate model integration. *J. Clim.* 19:6181–94
Fourier JBJ. 1827. Mémoire sur les températures du globe terrestre et des espaces planétaires. *Mem. Acad. R. Sci. Inst. Fr.* 7:569–604
Frame DJ, Booth BBB, Kettleborough JA, Stainforth DA, Gregory JM, et al. 2005. Constraining climate forecasts: the role of prior assumptions. *Geophys. Res. Lett.* 32:L09702
Geoffroy O, Saint-Martin D, Bellon G, Voldoire A, Olivié DJL, Tytéca S. 2013. Transient climate response in a two-layer energy-balance model. Part II. Representation of the efficacy of deep-ocean heat uptake and validation for CMIP5 AOGCMs. *J. Clim.* 26:1859–76
Gordon ND, Jonko AK, Forster PM, Shell KM. 2013. An observationally based constraint on the water-vapor feedback. *J. Geophys. Res. Atmos.* 118:12435–43
Gregory J, Forster P, Andrews T, Good P. 2016. The effective radiative forcing from explosive volcanic eruptions. *J. Clim.* Submitted
Gregory JM, Stouffer RJ, Raper SCB, Stott PA, Rayner NA. 2002. An observationally based estimate of the climate sensitivity. *J. Clim.* 15:3117–21
Hansen J, Sato M, Ruedy R, Kharecha P, Lacis A, et al. 2007. Climate simulations for 1880–2003 with GISS modelE. *Clim. Dyn.* 29:661–96
Hansen J, Sato M, Ruedy R, Nazarenko L, Lacis A, et al. 2005. Efficacy of climate forcings. *J. Geophys. Res.* 110:D18104
Hartmann DL, Klein Tank AMG, Rusticucci M, Alexander LV, Brönnimann S, et al. 2013. Observations: atmosphere and surface. See Stocker et al. 2013, pp. 159–254
Hegerl GC, Zwiers FW, Braconnot P, Gillett NP, Luo Y, et al. 2007. Understanding and attributing climate change. See Solomon et al. 2007, pp. 664–745
Johansson DJA, O'Neill BC, Tebaldi C, Haggstrom O. 2015. Equilibrium climate sensitivity in light of observations over the warming hiatus. *Nat. Clim. Change* 5:449–53
Karl TR, Arguez A, Huang B, Lawrimore JH, McMahon JR, et al. 2015. Possible artifacts of data biases in the recent global surface warming hiatus. *Science* 348:1469–72
Knutti R, Stocker TF, Joos F, Plattner GK. 2002. Constraints on radiative forcing and future climate change from observations and climate model ensembles. *Nature* 416:719–23
Koll DDB, Abbot DS. 2013. Why tropical sea surface temperature is insensitive to ocean heat transport changes. *J. Clim.* 26:6742–49
Kummer JR, Dessler AE. 2014. The impact of forcing efficacy on the equilibrium climate sensitivity. *Geophys. Res. Lett.* 41:3565–68
Levitus S, Antonov JI, Boyer TP, Baranova OK, Garcia HE, et al. 2012. World ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010. *Geophys. Res. Lett.* 39:L10603

Lewis N. 2013. An objective Bayesian improved approach for applying optimal fingerprint techniques to estimate climate sensitivity. *J. Clim.* 26:7414–29

Lewis N. 2015. Implications of recent multimodel attribution studies for climate sensitivity. *Clim. Dyn.* doi: 10.1007/s00382-015-2653-7

Lewis N, Curry J. 2014. The implications for climate sensitivity of AR5 forcing and heat uptake estimates. *Clim. Dyn.* 45:1009–23

Lin B, Chambers L, Stackhouse P Jr., Wielicki B, Hu Y, et al. 2010. Estimations of climate sensitivity based on top-of-atmosphere radiation imbalance. *Atmos. Chem. Phys.* 10:1923–30

Lindzen RS, Choi YS. 2009. On the determination of climate feedbacks from ERBE data. *Geophys. Res. Lett.* 36:L16705

Lindzen RS, Choi YS. 2011. On the observational determination of climate sensitivity and its implications. *Asia-Pac. J. Atmos. Sci.* 47:377–90

Loeb NG, Lyman JM, Johnson GC, Allan RP, Doelling DR, et al. 2012. Observed changes in top-of-the-atmosphere radiation and upper-ocean heating consistent within uncertainty. *Nat. Geosci.* 5:110–13

Loeb NG, Wielicki BA, Doelling DR, Smith GL, Keyes DF, et al. 2009. Toward optimal closure of the Earth's top-of-atmosphere radiation budget. *J. Clim.* 22:748–66

Manabe S, Wetherald RT. 1967. Thermal equilibrium of the atmosphere with a given distribution of relative humidity. *J. Atmos. Sci.* 24:241–59

Marvel K, Schmidt GA, Miller RL, Nazarenko LS. 2015. Implications for climate sensitivity from the response to individual forcings. *Nat. Clim. Change.* doi: 10.1038/nclimate2888

Masters T. 2014. Observational estimate of climate sensitivity from changes in the rate of ocean heat uptake and comparison to CMIP5 models. *Clim. Dyn.* 42:2173–81

Meraner K, Mauritsen T, Voigt A. 2013. Robust increase in equilibrium climate sensitivity under global warming. *Geophys. Res. Lett.* 40:5944–48

Merlis TM, Held IM, Stenchikov GL, Zeng F, Horowitz LW. 2014. Constraining transient climate sensitivity using coupled climate model simulations of volcanic eruptions. *J. Clim.* 27:7781–95

Morice CP, Kennedy JJ, Rayner NA, Jones PD. 2012. Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: the HadCRUT4 data set. *J. Geophys. Res.* 117:D08101

Murphy DM. 2010. Constraining climate sensitivity with linear fits to outgoing radiation. *Geophys. Res. Lett.* 37:L09704

Murphy DM, Forster PM. 2010. On the accuracy of deriving climate feedback parameters from correlations between surface temperature and outgoing radiation. *J. Clim.* 23:4983–88

Murphy DM, Solomon S, Portmann RW, Rosenlof KH, Forster PM, Wong T. 2009. An observationally based energy balance for the Earth since 1950. *J. Geophys. Res.* 114:D17107

Myhre G, Shindell D, Bréon FM, Collins W, Fuglestvedt J, et al. 2013. Anthropogenic and natural radiative forcing. See Stocker et al. 2013, pp. 659–740

Otto A, Otto FEL, Boucher O, Church J, Hegerl G, et al. 2013. Energy budget constraints on climate response. *Nat. Geosci.* 6:415–16

Ramanathan V, Cess RD, Harrison EF, Minnis P, Barkstrom BR, et al. 1989. Cloud-radiative forcing and climate: results from the Earth Radiation Budget Experiment. *Science* 243:57–63

Rhein M, Rintoul SR, Aoki S, Campos E, Chambers D, et al. 2013. Observations: ocean. See Stocker et al. 2013, pp. 255–316

Roe GH, Armour KC. 2011. How sensitive is climate sensitivity? *Geophys. Res. Lett.* 38:L14708

Rose BEJ, Armour KC, Battisti DS, Feldl N, Koll DDB. 2014. The dependence of transient climate sensitivity and radiative feedbacks on the spatial pattern of ocean heat uptake. *Geophys. Res. Lett.* 41:1071–78

Schwartz SE. 2007. Heat capacity, time constant, and sensitivity of Earth's climate system. *J. Geophys. Res.* 112:D24S05

Schwartz SE. 2012. Determination of Earth's transient and equilibrium climate sensitivities from observations over the twentieth century: strong dependence on assumed forcing. *Surv. Geophys.* 33:745–77

Sherwood SC, Bony S, Boucher O, Bretherton C, Forster PM, et al. 2015. Adjustments in the forcing feedback framework for understanding climate change. *Bull. Am. Meteorol. Soc.* 96:217–28

Sherwood SC, Bony S, Dufresne JL. 2014. Spread in model climate sensitivity traced to atmospheric convective mixing. *Nature* 505:37–42

Shindell DT. 2014. Inhomogeneous forcing and transient climate sensitivity. *Nat. Clim. Change* 4:274–77

Shindell DT, Faluvegi G, Rotstayn L, Milly G. 2015. Spatial patterns of radiative forcing and surface temperature response. *J. Geophys. Res. Atmos.* 120:5385–403

Skeie RB, Berntsen T, Aldrin M, Holden M, Myhre G. 2014. A lower and more constrained estimate of climate sensitivity using updated observations and detailed radiative forcing time series. *Earth Syst. Dyn.* 5:139–75

Soden BJ, Held IM. 2006. An assessment of climate feedbacks in coupled ocean-atmosphere models. *J. Clim.* 19:3354–60

Solomon S, Qin D, Manning M, Chen Z, Marquis M, et al., eds. 2007. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK: Cambridge Univ. Press

Spencer RW, Braswell WD. 2008. Potential biases in feedback diagnosis from observational data: a simple model demonstration. *J. Clim.* 21:5624–28

Stefan J. 1879. Über die Beziehung zwischen der Wärmestrahlung und der Temperatur. *Sitz. Math. Naturwiss. Cl. Kais. Akad. Wiss.* 1:391–428

Stevens B. 2015. Rethinking the lower bound on aerosol radiative forcing. *J. Clim.* 28:4794–819

Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, et al., eds. 2013. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK: Cambridge Univ. Press

Trenberth KE, Fasullo JT, O'Dell C, Wong T. 2010. Relationships between tropical sea surface temperature and top-of-atmosphere radiation. *Geophys. Res. Lett.* 37:L03702

Trenberth KE, Zhang Y, Fasullo JT, Taguchi S. 2015. Climate variability and relationships between top-of-atmosphere radiation and temperatures on Earth. *J. Geophys. Res. Atmos.* 120:3642–59

Tsushima Y, Abe-Ouchi A, Manabe S. 2005. Radiative damping of annual variation in global mean surface temperature: comparison between observed and simulated feedback. *Clim. Dyn.* 24:591–97

Tsushima Y, Manabe S. 2013. Assessment of radiative feedback in climate models using satellite observations of annual flux variation. *PNAS* 110:7568–73

Tyndall J. 1861. The Bakerian Lecture: on the absorption and radiation of heat by gases and vapours, and on the physical connexion of radiation, absorption, and conduction. *Philos. Trans. R. Soc.* 151:1–36

Wigley TML, Jones PD, Raper SCB. 1997. The observed global warming record: What does it tell us? *PNAS* 94:8314–20

Winton, M, Griffies SM, Samuels BL, Sarmiento JL, Frolicher T. 2013. Connecting changing ocean circulation with changing climate. *J. Clim.* 26:2268–78

Wong T, Wielicki BA, Lee RB, Smith GL, Bush KA, Willis JK. 2006. Reexamination of the observed decadal variability of the Earth radiation budget using altitude-corrected ERBE/ERBS Nonscanner WFOV data. *J. Clim.* 19:4028–40