The implications for climate sensitivity of AR5 forcing and heat uptake estimates

Nicholas Lewis[1] · **Judith A. Curry**[2]

Abstract: Energy budget estimates of equilibrium climate sensitivity (ECS) and transient climate response (TCR) are derived using the comprehensive 1750–2011 time series and the uncertainty ranges for forcing components provided in the IPCC Fifth Assessment Working Group I Report, along with its estimates of heat accumulation in the climate system. The resulting estimates are less dependent on global climate models and allow more realistically for forcing uncertainties than similar estimates based on forcings diagnosed from simulations by such models. Base and final periods are selected that have well matched volcanic activity and influence from internal variability. Using 1859–1882 for the base period and 1995–2011 for the final period, thus avoiding major volcanic activity, median estimates are derived for ECS of 1.64 K and for TCR of 1.33 K. ECS 17–83% and 5–95% uncertainty ranges are 1.25–2.45 K and 1.05–4.05 K; the corresponding TCR ranges are 1.05–1.80 K and 0.90–2.50 K. Results using alternative well-matched base and final periods provide similar best estimates but give wider uncertainty ranges, principally reflecting smaller changes in average forcing. Uncertainty in aerosol forcing is the dominant contribution to the ECS and TCR uncertainty ranges.

Keywords: climate sensitivity · transient climate response · energy budget · AR5

1. Introduction

The sensitivity of the Earth's climate to increasing concentrations of carbon dioxide (CO_2) is at the heart of the scientific debate on anthropogenic climate change. Climate sensitivity is a metric that is used to summarize the global surface temperature response to an externally imposed radiative forcing. The term 'equilibrium climate sensitivity' (ECS) refers to the equilibrium change in surface temperature to a doubling of atmospheric CO_2 concentration. A shorter-term measure of sensitivity, 'transient climate response' (TCR), represents the extent of global warming at the time of the CO_2 doubling following a linear increase in CO_2 forcing over a period of 70 years.

For three decades up to 2007, scientific assessments (including those by the Intergovernmental Panel on Climate Change – IPCC) provided a range and, generally, a best estimate for equilibrium climate sensitivity that hardly changed. In most cases the uncertainty range had a lower bound of 1.5 K and an upper bound of 4.5 K and the best estimate was 3 K. The 2007 IPCC Fourth Assessment Working Group I Report (AR4) narrowed the range to between 2.0 and 4.5 K, quantified as 'likely' (17–83% probability), influenced largely by climate model simulations. Subsequently, several observationally-based studies gave best estimates of between 1.5 K and 2K, substantially lower than most earlier studies. The IPCC Fifth Assessment Working Group I Report (AR5), published in 2014, reduced the 'likely' lower bound back to 1.5 K, making the range 1.5–4.5 K, reflecting the lower estimates that had been published recently in the literature. Significantly, the IPCC authors decided not to provide a best estimate for climate sensitivity in AR5.

The key issue faced in the AR5 assessment was interpreting the discrepancy between climate sensitivity estimates based on climate models (higher values) versus recent empirically-derived sensitivity analyses (lower values). A footnote to the AR5 Summary for Policy Makers (SPM) states: "No best estimate for equilibrium climate sensitivity can now be given because of a lack of agreement on values across assessed lines of evidence and studies."

[1] Bath, United Kingdom. email: nhlewis@btinternet.com

[2] School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia, USA.

AR5 considered estimates of ECS and TCR from various lines of evidence based on climate model simulations, recorded short- or long-term changes during the instrumental period since 1850 and temperature fluctuations as reconstructed from paleoclimate archives. AR5 stated that estimates based on ECS values of the CMIP5 atmosphere-ocean global climate models (AOGCMs) used in AR5 (the mean of which is 3.2 K) and analysis of feedbacks in models indicated an ECS range of 2 K–4.5 K. AR5 considered that paleoclimate ECS estimates based on past climate states very different from the present climate may differ from the sensitivity implied by the climate feedbacks of the present climate system. It assessed uncertainties in paleoclimate estimates of ECS as likely to be larger than for those from the instrumental record, and concluded [Section 10.8.2] that paleoclimate estimates only supported a wide uncertainty range (10-90%) of 1.0 - 6 K. AR5 [Section 12.5.3] also expressed doubts about ECS estimates based on timescales different from those relevant for climate stabilization (such as estimates based on climate response to volcanic forcing) or on forcings other than greenhouse gases (such as solar forcing). Accordingly, AR5 attributes its reduction in the bottom of the likely range for ECS to estimates using multidecadal data from the instrumental period, stating [Box 12.2] that the reduction "reflects the evidence from new studies of observed temperature change, using the extended records in atmosphere and ocean".

AR5 marginally reduced its estimate for TCR (relative to AR4), giving a likely range of 1.0 – 2.5 K, and an upper bound of 3 K at 95% probability. AR5 did not provide a best estimate for TCR, but the CMIP5 model mean TCR is approximately 1.8 K, similar to that for CMIP3 models.

Using a global energy budget approach, this paper seeks to understand the implications for climate sensitivity (both ECS and TCR) of the new estimates of radiative forcing and uncertainty therein given in AR5. This approach avoids to a substantial extent the dependence on AOGCM simulations in previous energy budget studies (e.g. Otto et al. 2013). Further, we refine the energy budget methodology for determining climate sensitivity to minimize the impact of natural internal variability on the estimate of climate sensitivity. And finally, we account carefully for the impact of uncertainties in forcing, ocean heat uptake and surface temperature on the determination of climate sensitivity.

The paper is structured as follows. The global energy budget approach is discussed in Section 2. Section 3 deals with data sources and uncertainties, Section 4 with choice of base and final periods, whilst methods are described in Section 5. Section 6 sets out the results, which are discussed in Section7.

2. Global energy budget approach

The basic global energy budget approach to determining climate sensitivity has been developed in Gregory et al (2002), Otto et al (2013) and Masters (2014). In this method, external estimates – observationally based so far as practical – of all forcing and Earth climate system heat uptake components, as well as of global mean surface temperature (GMST), are used to compute the mean changes in total forcing, ΔF, in total heat uptake, ΔQ, and in surface temperature, ΔT, between a base period and a final period. Equilibrium climate sensitivity (ECS) may then be estimated as:

$$\mathrm{ECS} = F_{2\times\mathrm{CO2}} \frac{\Delta T}{\Delta F - \Delta Q} \tag{1}$$

where $F_{2\times\mathrm{CO2}}$ is the radiative forcing attributable to a doubling of atmospheric CO_2 concentration. Total heat uptake by the Earth's climate system – the rate of increase in its heat content, very largely in the ocean – necessarily equals the net increase in energy flux to space (the Earth's radiative imbalance). The very simple model represented by equation (1) follows from conservation of energy.

Strictly, ECS as defined here is effective climate sensitivity: $\mathrm{ECS} = F_{2\times CO2} / \lambda$, where λ is the feedback parameter representing the net increase in energy flux to space per degree of surface warming given all feedbacks operating over the timescales involved, rather than equilibrium climate sensitivity, which requires the atmosphere-ocean system (but not ice sheets and other slow components of the climate system) to have reached a steady state. AR5 usually does not distinguish effective from equilibrium climate sensitivity, using equilibrium climate sensitivity and ECS to refer to estimates of both. However, it points out that in some climate models equilibrium climate sensitivity tends to be higher than the effective climate sensitivity (whether estimated from ΔF, ΔT and ΔQ or from other observations of transient climate change) because the feedbacks that are represented in the models (water vapour, lapse rate, albedo and clouds) vary with the climate state. The increase in model sensitivity may be linked to regional (Armour et al., 2013) and/or global (Meraner et al., 2013) rises in temperature. Moreover, even on a regional basis sensitivity may be affected by the forcing or ocean heat uptake pattern (Rose et al., 2014). These findings, which relate to specific models, depend on their latitudinal feedback patterns and cloud behaviour, which vary substantially between AOGCMs (Zelinka and Hartmann, 2012). It is unclear whether they apply to the real world.

As pointed out in AR5 (Bindoff et al., 2014, p.920), the transient climate response (TCR) represents a generic climate system property equalling the product of $F_{2\times CO2}$ and the ratio of the response of global surface temperature to a change in forcing taking place gradually over a ~70 year timescale. If most of the increase in forcing during a longer period occurs approximately linearly over the final ~70 years, then it likewise follows that:

$$\mathrm{TCR} = F_{2\times CO2} \frac{\Delta T}{\Delta F} \tag{2}$$

Both equations (1) and (2) assume constant linear feedbacks, and that ΔT is entirely externally forced . Otto et al. (2013) illustrated that the increase in total forcing over the last 70 years has approximated a linear ramp and constitutes most of the increase during the Instrumental period, implying that it is valid to estimate TCR using (2), provided that the final period is recent and the base period ends no later than about 1950. Numerical testing using a fitted 2-box constant-linear-feedbacks model (Fuglestvedt et al., 2008) confirms that, for the periods used here, applying (2) provides an estimate of TCR closely consistent with its formal definition.

Gregory and Forster (2008) also used an energy budget approach to estimating TCR, but based on linear regression rather than, as in the definition of TCR and here, comparing changes between two periods. Schwartz (2012) applied a similar method to several forcing datasets. A general energy budget framework has also been used elsewhere, for example in Armour and Roe (2011) and Roe and Armour (2011).

Although the energy budget approach does not use all of the available spatiotemporal observational data, the single equation model of the climate system involved follows directly from the conservation of energy, unlike more complex climate system models. Equations (1) and (2) do not assume a linear relationship between global changes in heat uptake and surface temperature, unlike the "kappa" model (Gregory and Forster, 2008). Whilst climate models are involved to an extent in forming the forcing estimates used and are needed for estimating Q for the base period, the energy budget method has greatly reduced dependence on complex climate models. This limited dependence on complex climate models makes energy budget estimates of ECS particularly robust with regard to dependence on model assumptions, subject to satisfactory forcing, heat uptake and surface temperature data being available.

Use in (1) and (2) of averages over base and final periods, rather than complete time series, captures much of the available information, since internal variability is high on sub-decadal timescales and only during the last decade or two has total forcing become reasonably large relative to its uncertainty. Base and final periods varying between one and four decades long have been used in energy budget studies (Gregory et al., 2002; Otto et al., 2013). There is a trade off in that longer periods reduce the effects of interannual and decadal internal variability (although not of multidecadal variability), but also reduce the magnitude of ΔF and make it difficult to avoid major volcanic activity.

Energy budget studies use only global mean data, which enables simpler and more robust quantification of uncertainties: uncertainty estimates for forcings often only relate to global mean values. The largest uncertainty is associated with aerosol forcing. Observationally-based climate sensitivity studies often form their own inverse estimates of aerosol forcing from recorded changes in temperature rather than, as in energy budget studies, using an external estimate that is not derived from temperature data. However, inverse estimates of aerosol forcing are unlikely to reflect as wide a range of uncertainties as does the AR5 estimate. Moreover, on a global scale, the estimated time evolutions of the magnitudes of aerosol and greenhouse gas forcings are too highly correlated (for the AR5 best estimates, $r = 0.98$ over 1850–2011) for reliable separate estimation. Therefore, inverse estimates of aerosol forcing require a model and data that resolve surface temperature by hemisphere (aerosol forcing being concentrated in the northern hemisphere), at least.

The energy budget approach was attempted over a decade ago (Gregory et al, 2002), but at that time the estimated change in forcing was relatively small and poorly constrained, and ocean heat content data was poor. The result was a very high median ECS estimate of 6.1 K, with a long tail on the estimated probability distribution extending up to infinity, and beyond to negative values (corresponding to the possibility of the divisor $\Delta F - \Delta Q$ in (1) being negative).

Recently, Otto et al (2013) used an energy budget approach to estimate both ECS and TCR, estimating ΔF and $F_{2\times CO2}$ from simulations by a large sample of CMIP5 AOGCMs, with an adjustment made to reflect the high level of aerosol forcing in CMIP5 models. This yielded encouragingly stable estimates for ECS using final periods of 1980-89, 1990-99, 2000-09 and 1970-2009, with medians of 2.0 K when using 2000-09 data and 1.9 K otherwise. However, it is unclear whether the spread of estimates of ΔF from a sample of AOGCMs is representative of the true uncertainty therein, even if their median is adjusted to be in line with a best estimate that is more observationally based. And whilst the estimate using a final period of 2000–09 (when radiative forcing was highest and volcanic activity low) should in principle provide the most reliable estimate of ECS, estimates of the climate system's heat uptake over that period vary considerably. Had Otto et al (2013) replaced its 2000–09 heat uptake estimate of 0.73 Wm^{-2} with that of 0.5 Wm^{-2} from Loeb et al (2012), its ECS estimate based on that period would have been 1.7 K rather than 2.0 K.

The traditional concept of radiative forcing is stratospherically-adjusted radiative forcing (RF), computed with all tropospheric properties held fixed at their unperturbed values. That is not a suitable measure for an energy budget approach, since the feedback parameter derived with respect to RF can vary significantly for different forcing agents (Forster et al., 2007). Hansen et al. (2005) introduced the term efficacy for the surface temperature response to RF from a particular agent relative to the response to RF from CO_2, and it is efficacy-adjusted RF that is appropriate for an energy budget approach. The Forster et al. (2013) radiative forcing estimates used in Otto et al. (2013) were diagnosed for CMIP5 models from their surface temperature responses, and therefore reflect forcing efficacies. The concept of effective radiative forcing (ERF) introduced in AR5 includes many of the rapid adjustments that differ among forcing agents, thereby reflecting in most cases their relative efficacy. References here to forcing (including to $F_{2\times CO2}$) should be understood as relating to ERF, except where otherwise indicated. The relationship of ERF to RF can only be estimated using AOGCMs; for some forcing agents estimates of

RF also have to be derived using AOGCMs. Whilst the AR5 ERF estimates are supported by multiple studies and thorough assessment, expert judgement also played a role. AR5's best estimate of ERF differs from that of RF only for aerosols, but its uncertainty ranges are generally wider for ERF than RF.

The main source of uncertainty in estimating both ECS and TCR using an energy budget approach is uncertainty in ΔF . This uncertainty accounts for most of the asymmetry in probability distributions of observationally-based estimates for TCR and, particularly, ECS (Roe and Armour, 2011). Fractional uncertainty in ΔT is substantially smaller, even allowing for internal variability. Although uncertainty in ocean heat content remains significant, it has been reduced since the introduction of the Argo network (Lyman and Johnson, 2014). Progress made in resolving biases in historical measurements has also helped (Willis et al, 2009). Moreover, ΔQ is only 20-30% of ΔF for the periods used in this study, so fractional uncertainty in ΔQ contributes much less to uncertainty in $\Delta F - \Delta Q$ than does the same fractional uncertainty in ΔF . The two main contributors to uncertainty in ΔF are aerosols and, to a substantially smaller extent, well-mixed greenhouse gases (WMGG), principally CO_2. Uncertainty in forcing from WMGG almost entirely relates to how much forcing a given concentration of each greenhouse gas produces; uncertainty in concentrations is minor. Following Otto et al. (2013), forcing uncertainty is taken to be strongly correlated between CO_2 and other WMGG. Since $F_{2\times CO2}$ appears in the numerator of (1) and (2), and ΔF (of which WMGG forcing is by far the largest component) in the denominator, the effects on ΔF and on $F_{2\times CO2}$ of uncertainty in forcing from WMGG cancel out to a substantial extent. As a result, despite using twice as high uncertainty in WMGG forcing here as in Otto et al. (2013) – ±20% rather than ±10% – the largest contributor by far to uncertainty in ECS and TCR estimation is aerosol forcing.

3. Data sources and uncertainties

Data and uncertainty estimates identical to those given in AR5 WGI have been used unless stated otherwise. Forcing, heat uptake and temperature data are considered in turn.

3.1 Forcings

ERF time series medians are sourced from Table AII.1.2 of AR5. Uncertainty estimates for 2011 are primarily taken from Table 8.6 of AR5. Where no separate ERF estimate is given the treatment in Table 8.SM.5 is followed, with RF being treated as representative of ERF but with an additional uncertainty of ±17% of the RF added in quadrature. AR5 states that black carbon (BC) on snow forcing ($ERF_{BC\ on\ snow}$, estimated at +0.04 Wm^{-2} in 2011) causes a 2–4 times larger GMST change per unit forcing than does CO_2. BC on snow ERF is therefore efficacy-adjusted, estimated by scaling the Table 8.6 uncertainty range distribution by a 2–4 efficacy 5–95% range.

Following AR5, symmetrical 5–95% ranges are taken to represent independent Gaussian distributions. For asymmetrical uncertainty ranges – those not symmetrical about the central (median) estimate –.the 5%, 50% and 95% points are all shifted equally by the amount needed to obtain a (unique) fit to a lognormal distribution for the shifted variable. A fitted shifted lognormal distribution (Meinshausen, 2009) is a natural way of representing a skewed mono-peaked smooth distribution when only three percentile points are given. It generalises the lognormal distribution to provide, according to the shift magnitude, anywhere between a symmetrical normal distribution and a highly skewed lognormal distribution. As a cross-check, ten million samples were drawn from each of the calculated anthropogenic ERF distributions. The computed distribution of their sum accurately matched the probability density function shown in AR5 Figure 8.16 and the 5%, 50% and 95% percentiles of 1.13, 2.29 and 3.33 Wm^{-2} given in AR5.

The time evolutions of all significant constituents of anthropogenic forcing are highly correlated. Over 1850-2011, the absolute correlations of the time series in AR5 Table AII.1.2 between ERF for WMGG (ERF_{WMGG}) and ERF for the other significant anthropogenic forcing components (Ozone, Aerosol and Land use change) are all 0.95 or higher, and the correlation of 'GHG Other' ERF with CO_2 ERF exceeds 0.99. Moreover, as discussed in Section 2, uncertainty in forcing from WMGG almost entirely relates to how much radiative forcing a given concentration of each greenhouse gas produces and is likely strongly correlated among the WMGG. Reflecting these considerations, and the need to sample uncertainty in $F_{2\times CO2}$, Otto et al. (2013) additively combined forcings (adding their uncertainties in quadrature) into three component time series: natural (sum of solar and volcanic), WMGG and other anthropogenic (residual).

Otto et al. then sampled current values for the two anthropogenic forcing components and scaled the samples to match the forcing best estimate time series. This treatment assumes both that uncertainty is fractional (proportional to the magnitude of the forcing) in nature and that it affects the base and final periods identically. Uncertainty in natural forcing was not scaled. Otto et al also used the uncertainty realizations provided by the WMGG forcing samples to scale the central $F_{2\times CO2}$ value in the same proportions as the central WMGG forcing value were scaled, hence providing $F_{2\times CO2}$ samples with uncertainty realizations (proportionately) matching those for WMGG forcing. Adding in each period the mean forcing components and then subtracting total forcing in the base period from that in the final period gave sample values for ΔF, which were matched with corresponding sample values of $F_{2\times CO2}$. A similar approach is used here, but with forcing split into seven rather than three components and a more detailed treatment of uncertainties, that, in particular, allows where appropriate part of fractional uncertainty in a forcing component to be independent between the base and final periods and for (similarly independent) fixed uncertainty elements to exist as well. Further information is set out under 'Methods'.

Volcanic forcing ($ERF_{Volcano}$) invokes a complex radiative-dynamical response, and it is doubtful that on an ERF basis its effect on global surface temperature is comparable to that for other forcings. Tomassini et al (2007) found an optimum scaling factor of 0.6 for volcanic forcing in order to best fit the global temperature response to their simple climate model. This means that 1 Wm^{-2} of volcanic forcing had a comparable effect on GMST to 0.6 Wm^{-2} of CO_2 forcing, which implies that volcanic forcing needs to be multiplied by 0.6 before combining the various forcing components. Ring et al. (2012) found the same. Meinshausen et al (2011) also found that they needed to apply a scaling factor of less than one to volcanic forcing. For major eruptions, their scaled estimates are 50-65% of the AR5 volcanic forcing estimates, which are more negative than some previous estimates: they average 15% more negative for 1880–2011 relative to the GISS dataset (http://data.giss.nasa.gov/modelforce/Fe.1880-2011.txt). This suggests applying a scaling factor of 0.5–0.6 rather than 0.6 as found when using previous volcanic forcing estimates. A scaling factor of 0.5–0.6 does indeed make ECS and TCR estimates least sensitive to differences in volcanic forcing between the base and final periods. Nevertheless, since no efficacy range for volcanic forcing is given in AR5, it is taken as unity for the main results. The issue is instead addressed by using base and final periods with matching mean estimated volcanic forcing. However, when investigating the sensitivity of estimates to the choice of base and final periods involves a significant mismatch in their mean volcanic forcing, the results of applying an efficacy scaling factor of 0.55 are shown.

For all the base and final periods considered, total solar irradiance (TSI) forcing estimates in AR5 (which provide its ERF estimate – ERF_{Solar}) have means matching within ~0.05 Wm^{-2}, so even if TSI had an efficacy considerably different from unity the resulting bias would be modest. AR5 suggests that variability in aspects of solar activity other than TSI, in particular in the ultraviolet component of

TSI, may have an enhanced climate impact. However, since no quantification is given, that possibility is ignored here.

There are arguments for making two adjustments to the AR5 forcing estimates. One is to adjust the 2011 aerosol ERF ($ERF_{Aerosol}$) expert judgement best estimate of −0.9 Wm^{-2} by +0.12 Wm^{-2}, to the mean (−0.78 Wm^{-2}) of the estimates from the six satellite-based studies that were used in AR5, thus forming estimates of ECS and TCR that are more independent of global climate models. The other is to adjust land use change (LUC) forcing (ERF_{LUC}). AR5 assesses the surface albedo effect of LUC as an RF and ERF of −0.15 Wm^{-2} in 2011. At the same time, AR5 concludes that, accounting for other processes, land use change is about as likely as not to have caused a net cooling of the Earth's surface, implying a best estimate of zero ERF_{LUC}. Since the intention here is to show the implications of AR5 best estimates of forcings and energy inventory, for the main results no adjustment to either of these forcing estimates is made. However, sensitivity to setting ERF_{LUC} to zero is shown; shifting the 2011 $ERF_{Aerosol}$ distribution by +0.12 Wm^{-2} has almost the same effect.

The forcing best estimates and uncertainties used for the main results are summarised in Table 1.

ERF component	AR5 1750–2011 best estimate and 90% CI	Fractional-type uncertainty used	Part treated as independent	Added fixed uncertainty
WMGG	2.831 ± 0.57	± 0.57	0%	
Ozone (total)	*0.350 ± 0.209*			
Stratospheric H_2O	*0.073 ± 0.051*			
Land use (sfc albedo)	*-0.150 ± 0.103*			
Total nonGABC	0.273 ± 0.239	± 0.239	50%	
Aerosol (total)	-0.90 (-1.00, +0.80)	-1.00, +0.80	25%	
BC on snow	0.040 (-0.021, +0.050)	-0.021, +0.050	Ignored	
Contrails	0.05 (-0.03, +0.10)	-0.03, +0.10	Ignored	
Total anthropogenic	*2.294(-1.16, +1.04)*			
Solar	0.03 ± 0.05	± 0.05	50%	±0.05
Volcanic	-0.125 ± 0.035	± 0.035	50%	±0.072

Table 1 Components of ERF and treatment of their uncertainties. Units are Wm^{-2}

3.2 Total heat uptake

Estimation of final period mean heat uptake is derived from the climate system energy accumulation observational best estimates and uncertainty ranges shown in Box 3.1, Figure 1 of AR5. These estimates include 0–700 m OHC from an update of Domingues (2008), 700–2000 m ocean heat content (OHC) from Levitus et al. (2012), and allowances for minor heat uptake by the abyssal (2000–6000 m) ocean, ice melt, land and atmosphere, as described in detail in Box 3.1 of AR5. The ocean accounts for over 90% of total estimated energy accumulation and for almost all uncertainty. The AR5 OHC estimates are three year (0–700 m) and five year (700–2000 m) running means, which reduces noise. Heat uptake is

estimated from the annualised difference in accumulated energy between the start and end years in each period, with all uncertainties being added in quadrature. Estimates based on regression slope are similar.

Since no usable OHC measurements exist before about 1950 and independent sea level rise estimates are insufficiently precise, estimates derived from AOGCM simulations are used. Gregory et al. (2013) provides graphically a time series of global mean sea level (GMSL) rise due to thermal expansion from 1860 based on a simulation by the CCSM4 AOGCM, with volcanic forcing included, starting in 850. The estimated rises in GMSL over 1860–1882, 1860–1900 and 1930– 1950 amount to mean rates of respectively 0.55, 0.35 and 0.71 mm yr^{-1}. Church et al. (2011, corrected 2013) provide estimates for the rise in both steric GMSL and total OHC over 1972–2008 and 1993–2008, from which the rate of ocean heat uptake per unit area of the Earth's surface corresponding a 1 mm yr^{-1} rise in GMSL can be derived. The implied GMSL-to-ocean-heat-uptake conversion rates for the 1972–2008 and 1993–2008 periods are within 10% of each other, and their mean is within 10% of the estimate based on 1955–2010 data for the 0–2000 m ocean layer in Levitus et al. (2012). Taking their average of 0.47 Wm^{-2} per 1 mm yr^{-1}, the Gregory et al. (2013) GMSL rise rates equate to 0.26 Wm^{-2} over 1860–1882, 0.16 Wm^{-2} over 1860–1900 and 0.33 Wm^{-2} over 1930–1950. The estimate for 1860–1900 agrees with that in Gregory et al. (2002) using a different AOGCM. According to AR5 estimates, mean total forcing over 1850–1900 was almost the same as that over 1860–1900, whilst forcing over 1859–1882 was marginally lower than over 1860–1882. However, the CCSM4 model has TCR and ECS values of 1.8 K and circa 3.0 K that are some 35–85% higher than the best estimates for those parameters arrived at in this study. We therefore take only 60% of the base period heat uptake estimated from the Gregory et al. (2013) simulations, giving 0.15 Wm^{-2} for 1859–1882, 0.10 Wm^{-2} for 1850–1900 and 0.20 Wm^{-2} for 1930–1950. Heat uptake by other components of the climate system is ignored. Uncertainty in these estimates is difficult to quantify. Gregory et al. (2013) gives a 0.2 mm yr^{-1} range over the last 150 years, equating to ±0.03 Wm^{-2} after scaling down to 60%. This is consistent with the unscaled ±0.04 Wm^{-2} range for 1861–1900 given in Gregory et al. (2002). The implied standard error of 0.02 Wm^{-2} seems unrealistically low; we use 50% of the OHU estimate.

The heat uptake best estimates and uncertainties used for each period are summarised in Table 2.

Period	Estimated rate of total heat uptake	Standard error used	Source of estimate and error used
1859–1882	0.15	0.075	Gregory et al. (2013), scaled & higher uncertainty
1850–1900	0.10	0.05	As above
1930–1950	0.20	0.10	As above
1971–2011	0.43	0.073	AR5: all heat uptake components and uncertainties
1987–2011	0.51	0.107	As above
1995–2011	0.51	0.087	As above

Table 2 Heat uptake estimates with their uncertainties and sources. Units Wm^{-2}

Internal OHC trend variability, used as a proxy for variability in total heat uptake, is estimated from years 2,100 to 6,100 of the long HadCM3 (Gordon et al., 2000) control run. Changes have a standard deviation equating to 0.045 Wm^{-2} over the Earth's surface for a period of 16 years, the shortest used here. The same heat uptake variability is used for longer periods.

Since ΔQ and ΔF are positive, estimates of TCR are better constrained than estimates of ECS, due both to uncertainty in ΔQ being absent and to the uncertainty in ΔF being applied to ΔF rather than to $\Delta F - \Delta Q$. In principle, estimation of ECS using an energy budget approach should be little affected by variability in ocean heat uptake, although estimation of TCR might be affected by this variability. Given that ECS and $F_{2\times CO2}$ are assumed to be fixed albeit uncertain parameters, (1) implies that, for fixed ΔF, fluctuations in ΔQ must be offset by opposing fluctuations in ΔT otherwise energy would not be conserved globally. Decadal internal variability should only affect ECS estimation to the extent that it changes total forcing, not just energy interchange between the ocean and atmosphere. If decadal internal variability increases OHC, energy conservation implies that the surface temperature should be depressed. There is indeed negative correlation in the HadCM3 control run between mean surface temperature and change in OHC, but it is small over 15–25 year periods and is ignored.

3.3 Surface temperature

Of the three GMST datasets cited in AR5, only HadCRUT4 (Morice et al, 2012) covers the 1859–1882 and 1850–1900 base periods used in this study; it is therefore employed. The other two GMST datasets cited in AR5 commence in 1880, and from 1880–1900 to each of the final periods used in this study both datasets show almost identical increases in GMST to HadCRUT4's increase. HadCRUT4 also has the advantages of including corrections for a change in the way sea surface temperatures were measured in the mid-twentieth century and of providing an ensemble of 100 temperature realizations that preserves the time-dependent correlation structure. Uncertainty in GMST for each period is calculated on a basis consistent with the applicable covariance matrix of observational uncertainty.

The standard deviation of internal (unforced) variability in ΔT is estimated at 0.06 K, from the standard deviation of the difference in the means of 16 year periods centred 128 years apart in years 2,100 to 6,100 of the HadCM3 control run. For a 24 year period, the estimate drops to 0.05 K. Reducing the separation to 64 or 32 years has little effect. There is no evidence from comparisons of the surface temperature power spectrum of the HadCM3 control run with 1901–2010 observations and last millennium reconstructions (AR5 Figure 9.33) that the HadCM3 control run is lacking in interannual, decadal or multidecadal internal variability. Moreover, care has been taken to match the influence on the base and final periods of the most obvious source of quasi-regular multidecadal internal variability, the Atlantic multidecadal oscillation (AMO). Therefore, the impact of such variability should be considerably lower than with randomly selected periods.

However, global temperature records are not long enough to fully quantify multidecadal internal variability. AR5 notes (Hartmann et al., 2014, p.230) that quantifying spatio-temporal climate variability and its trends is hampered by relative shortness of and uncertainties in climate records and the presence of large variability on decadal and multidecadal timescales. Recognising this, we take a standard deviation for internal variability in ΔT of 0.08 K, which equates to adding in quadrature almost the same amount of variability again to the level derived from the HadCM3 control run. We then similarly combine that standard deviation with the GMST observational uncertainties for the base and final periods. Furthermore, we investigate sensitivity to making a further 50% increase in internal variability in ΔT.

4. Choice of base and final periods

A key decision in an energy budget study is the choice of base and final periods. Measurement uncertainties, particularly for ocean heat content, and internal variability, make periods significantly shorter than a decade unsuitable. The main consideration is then obtaining a large value of $\Delta F - \Delta Q$

(just ΔF for TCR estimation) relative to uncertainty therein, since fractional uncertainty in ΔT is smaller. The final period should end as recently as possible, since forcing is highest then and uncertainties generally smallest. The AR5 forcing estimates are for 1750–2011 and its heat content data ends in 2011, hence 2011 is the obvious end date. Whilst using a final period several decades long has some attractions, as in Gregory et al (2002), there are several disadvantages to doing so. One is that both $\Delta F - \Delta Q$ and ΔT, and their respective ratios to uncertainty therein, decline with the length of the final period. But another important issue is uncertainty as to the magnitude and efficacy of volcanic forcing.

The available datasets exhibit disagreement as to mean volcanic aerosol optical depth that is substantial for recent eruptions, and very large for earlier eruptions (Myhre et al., 2014). Moreover, as discussed in Section 3.1, although AR5 does not quantify the efficacy of volcanic forcing several studies point to it being substantially below one. In view of the major uncertainties as to the magnitude and efficacy of volcanic forcing it is highly desirable to reduce the incidence of major volcanic episodes in both the base and the final periods. So far as the final period is concerned, that points to it extending no further back than 1995, as forcing from the 1991 eruption of Mount Pinatubo was substantial (taken as <-0.5 Wm^{-2}) through to 1994. The longest available early base period without substantial volcanic forcing is 1859–1882, which per the AR5 estimates has almost the same low average level of volcanic forcing (~ -0.1 Wm^{-2}) as 1995–2011. The better observed period 1930–1950, which has similarly low volcanic activity but higher total forcing, is also used as a base period. The long base period 1850–1900, used in AR5 to represent pre-industrial GMST, has much greater average volcanic forcing, estimated as -0.37 Wm^{-2} in AR5, so it is not well matched with 1995–2011.

Apart from avoiding high levels of volcanic forcing and the associated uncertainties, there are two other principal considerations in choosing base and final periods. One is minimising uncertainty in ΔQ. For a final period ending in 2011, the energy accumulation uncertainty estimates shown in Box 3.1, Figure 1 of AR5 imply that uncertainty in mean heat uptake rate does not vary greatly with start year from 1971 to 1995, due to the compensating effects of uncertainty in starting energy content and the length of period. But there is an additional consideration, in that the AR5 energy estimates and their uncertainty reflect only one of several datasets of 0–700 m layer OHC – an update of Domingues et al. (2008). Data corresponding to 0–700 m OHC estimates from Levitus et al. (2012), Ishii and Kimoto (2009), Smith and Murphy (2007) and (covering 1993 onwards) Lyman et al. (2010), all updated to 2011, were also downloaded. The AR5 0–700 m OHC change-to-2011 estimate agrees much better with the average of estimates from these other datasets over some periods than others. Agreement is close (the difference between the average of the ocean heat uptake rate (OHU) estimates from the other datasets and from the updated Domingues 2008 dataset is under 5% of $\Delta F - \Delta Q$) with starting years in the ranges 1981–1983, 1987–1995 and 1997–1998. That supports choosing the final period as 1995–2011. For testing sensitivity to the choice of a different, longer period the clear second choice is 1987–2011: using an early 1980s starting year results in significant extra volcanic ERF uncertainty, from El Chichon's eruption, for a relatively small increase in period length. Since 1987– 2011 has lower average forcing than 1995–2011, it provides less well-constrained sensitivity estimates. We give results using both periods, and also 1971–2011, the period starting in the 1970s with the lowest volcanic forcing; 1971–2011 has only modest disagreement (6%) in OHU between the Domingues dataset and the average of the other datasets. Both 1987–2011 and 1971–2011 have average volcanic forcing that closely matches that in 1850–1900.

As an alternative to selecting base and final periods with low and well-matched volcanic activity, another way of minimising the uncertainties involved in volcanic forcing is to carry out a regression analysis on a time series filtered to exclude years with significant volcanic forcing (Gregory and Forster, 2008). Adopting their volcanic forcing threshold of -0.5 Wm^{-2}, and regressing GMST on total forcing over 1850–2011, would produce almost the same estimate for TCR as that using (as here) the two-period method with an 1859–1882 base period and 1995–2011 final period.

The other principal consideration in choice of base and final periods is matching between them the effects of interannual and multidecadal internal variability. El Niño/Southern Oscillation (ENSO) is the most important coupled ocean-atmosphere phenomenon to cause global climate variability on interannual time scales. The MEI (Wolter and Timlin, 1993) and extended MEI (MEI.ext: Wolter and Timlin, 2011) indices provide estimates of the ENSO state respectively from 1950 onwards and from 1871–2005. They show that 1995–2011 and 1930–1950 have closely matching mean ENSO states, in terms of their average MEI (and comparable MEI.ext) index values. However, the average MEI value for all final periods to 2011 starting before 2005 exceeds the comparable MEI.ext value for all two decade or longer periods within 1850–1930 (before 1871, treating MEI.ext as always having its average value). Nevertheless, for 1859–1882 and 1850–1900 base periods the difference in average index value compared with the final period is only 10% (for 1995–2011) or 15% (for 1987–2011 and 1971-2011) of the MEI change associated with the exceptional 1997/98 ENSO. These differences, while likely to bias ΔT – and hence estimates of ECS and TCR – slightly upwards, are well within natural variability.

Multidecadal variability is a greater concern, since it will be highly correlated within each of the base and the final periods. Tung and Zhou (2013) identify a ~0.2 K peak-to-peak quasi-cyclical fluctuation in GMST that is associated with the AMO, with approximately a 65 year quasi-periodicity, peaking around 1875, 1940 and 2005. The Delsole et al. (2011) findings relating to sea surface temperature are consistent with this. Examination suggests 1859–1882 and 1930–1950 match better with 1995–2011, and 1850–1900 with 1987–2011, for multidecadal fluctuations than other combinations. All combinations are reasonably matched except those involving 1971–2011, which appears less positively affected by the AMO than any of the base periods. It has been paired with the 1850–1900 base period, which involves the least mismatch.

Figure 1 shows variations in the three sources of natural variability that have been discussed. Five year running means are shown for the MEI.ext index and AMO index (Enfield et al, 2001) since annual fluctuations are irrelevant. Annual data is shown for volcanic forcing since period start and end years are selected to avoid volcanic episodes.

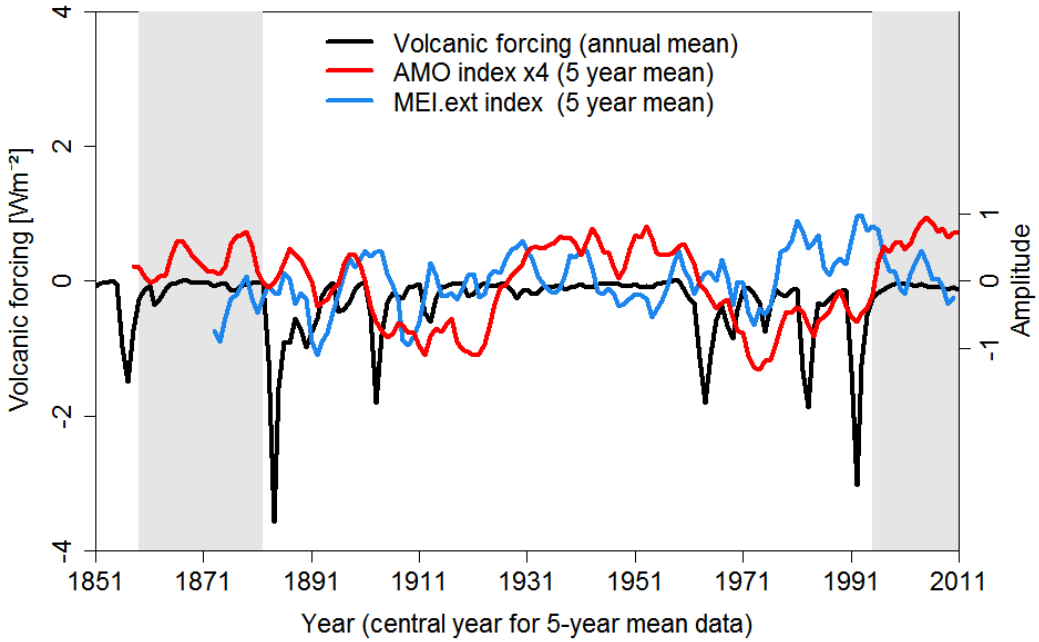

Fig. 1 Natural factors that influence selection of base and final periods. Volcanic forcing is from AR5. The AMO index has been scaled up by 4 times. Post-2005 MEI.ext index values have been derived from MEI index values. The units of the two indices are arbitrary. The two epochs chosen as the best-case base and final periods are shaded.

5. Methods

The method used is essentially identical to that in Otto et al (2013), save for a more sophisticated treatment of forcing uncertainties. The main steps in deriving best estimates and uncertainty ranges for ECS and TCR for each base period and final period combination are as follows:

1) Derive, as described in Section 3.1, 2011 distributions for fractional uncertainty for all forcings, combining CO_2 and GHG Other forcings into ERF_{WMGG}, and ozone (tropospheric and stratospheric), land use change and stratospheric water vapour forcings into $ERF_{nonGABC}$. The three forcings with uncertainty ranges in AR5 that are not symmetrical about their best (median) estimates – $ERF_{Aerosol}$, $ERF_{BC\ on\ snow}$ and combined contrails and contrail-induced cirrus ("Contrails") forcing ($ERF_{Contrails}$) – are treated as separate components, since asymmetrical uncertainty ranges cannot simply be added in quadrature. The foregoing treatment of AR5 forcings and their uncertainties results in separate 2011 uncertainty distributions for seven components making up total forcing: ERF_{WMGG}, $ERF_{Aerosol}$, $ERF_{BC\ on\ snow}$, $ERF_{Contrails}$, $ERF_{nonGABC}$, $ERF_{Volcano}$ and ERF_{Solar}. Where uncertainty in an ERF component is treated as partially independent between the base and final periods, the total 2011 uncertainty variance is split appropriately between separate common and independent random elements, such that when samples of each of these elements are added the original 2011 median and uncertainty range for the ERF component are achieved. Similarly, any fixed element in the total 2011 uncertainty variance (represented as a zero-mean Gaussian distribution) is separated out.

2) Compute uncertainty distributions for ΔT (using the HadCRUT4 ensemble of 100 temperature realizations) and for ΔQ, adding in quadrature the estimated uncertainties for the base and final period means, and for internal variability.

3) Draw from their uncertainty distributions a large number of random samples of ΔT, ΔQ and the 2011 ERF values of each of the (in some cases combined) forcing components. Sample separately for the base and final periods period fractional uncertainty elements that are independent between base and final periods. Also where relevant sample separately for the base and final periods fixed elements of uncertainty in forcing components.

4) Derive ΔF by scaling the 2011 ERF samples to match their best estimate time series, aggregating the results and taking the difference in the period means of the sum of the aggregated values between the base and final periods (using in the base and final periods the separate samples of fixed and the independent fractional uncertainty elements). Create samples of $F_{2\times CO2}$ by scaling the sampled 2011 WMGG ERF values.

5) For each sample realization of ΔT, ΔF, ΔQ and $F_{2\times CO2}$, compute the ECS and TCR value given by equations (1) and (2).

6) Compute a fine-binned histogram of the sample ECS and TCR values to give their best (median) estimates and uncertainty ranges, treating values where the denominator in (1) or (2) is negative as infinitely high.

Further details are now given of each of the foregoing steps.

5.1 Deriving forcing uncertainty distributions

Part of the fractional uncertainty in the residual anthropogenic component ($ERF_{nonGABC}$) – that remaining after excluding $ERF_{Aerosol}$, $ERF_{BC\ on\ snow}$, $ERF_{Contrails}$ and ERF_{WMGG} – is treated as independent between the base and final periods. That is, for $ERF_{nonGABC}$ separate random uncertainty realizations apply to the two periods in respect of the independent part of the assumed total fractional uncertainty.

Base period	Final period	ΔT [K]	ΔF [W m^{-2}]	ΔQ [W m^{-2}]
1859–1882	1995–2011	0.71 (0.56–0.86)	1.98 (0.99–2.86)	0.36 (0.15–0.58)
1850–1900	1987–2011	0.66 (0.52–0.81)	1.88 (0.92–2.74)	0.41 (0.19–0.63)
1850–1900	1971–2011	0.52 (0.38–0.66)	1.57 (0.69–2.36)	0.34 (0.16–0.52)
1930–1950	1995–2011	0.49 (0.35–0.63)	1.38 (0.53–2.15)	0.31 (0.07–0.56)
Otto et al (2013) estimates for comparison				
1860–1879	*2000–2009*	*0.75 (0.55–0.95)*	*1.95 (1.37–2.53)*	*0.65 (0.38–0.92)*
1860–1879	*1970–2009*	*0.48 (0.28–0.68)*	*1.21 (0.69–1.73)*	*0.35 (0.22–0.48)*

Table 3 Best estimates (medians) and 5–95% uncertainty ranges for changes ΔT in Global Mean Surface Temperature, ΔF in Effective Radiative Forcing and ΔQ in total climate system heat uptake between the base and final periods indicated. The final two lines, in italics, show comparative values for Otto et al (2013) for the periods highlighted in that paper and used in AR5

The remaining (common) part of fractional uncertainty is derived from another separate, single, random uncertainty realization which is applied in both periods. The common parts of fractional uncertainty therefore only affect ΔF to the extent that forcing magnitudes (with which fractional uncertainty scale) differ between the periods. The same treatment applies to volcanic forcing ($ERF_{Volcano}$), solar forcing (ERF_{Solar}) and $ERF_{Aerosol}$. This treatment, more cautious than that in Otto et al. (2013), is appropriate where a substantial proportion of the uncertainty relates to the level of the factor causing the forcing (such as emissions of ozone precursors) rather than (as for WMGG) to uncertainty as to the ERF resulting from a known level of the factor.

Uncertainties in $ERF_{BC\ on\ snow}$ and $ERF_{Contrails}$ are small: for simplicity, no part of either is treated as independent between base and final periods. For $ERF_{nonGABC}$, 50% of the total variance is taken to be independent between base and final periods, since uncertainty therein is dominated by uncertainty in tropospheric ozone forcing, which Stevenson et al. (2013) suggest is fairly evenly split between the two types of uncertainty (level of the factor causing the forcing, and the ERF per unit of that factor). The independent part is also taken as 50% for both solar and volcanic forcing; results are insensitive to these choices.

The independence issue is most important for $ERF_{Aerosol}$, which not only has the largest 2011 uncertainty but has a magnitude that is non-negligible in the base period. Smith et al (2011) indicate that uncertainty in sulphate emissions is relatively small, whilst Bond et al. (2013) indicate the opposite for black carbon and related aerosols. Carslaw et al. (2013) find most uncertainty in indirect aerosol forcing to relate to emissions rather than aerosol processes, with uncertainty in natural emissions in 1750 contributing more than uncertainty in anthropogenic emissions due to the strongly nonlinear relationship of indirect aerosol forcing to total emissions. However, their 5%-to-95% uncertainty is only 0.72 Wm^{-2}, much smaller AR5's 1.80 Wm^{-2} range for total aerosol forcing, and accounts (based on those ranges) for only some 16% the total AR5 $ERF_{Aerosol}$ variance. Emission uncertainty is also lower for the multiyear periods used here than for individual years. Moreover, Carslaw et al. use a subjective Bayesian statistical approach, which may give unrealistic uncertainty estimation when (as with aerosol forcing) strongly non-linear functional relationships are involved (Lewis, 2013). In AR5, Boucher, Randall et al. (2014) discuss various sources of uncertainty in aerosol forcing but do not quantify their separate contributions to the overall assessed uncertainty. However, the large discrepancy between model-based

and satellite-based $ERF_{Aerosol}$ estimates suggests that much of the AR5 uncertainty range relates to poor understanding of aerosol processes and, particularly, of related cloud processes and properties (Zelinka et al, 2014). There are also other non-emission uncertainties: in radiative properties of aerosols, in surface albedo, etc. Moreover, uncertainty in 1750 aerosol forcing is fully allowed for in AR5's 2011 $ERF_{Aerosol}$ range. This all points to only a minority of the base period uncertainty in $ERF_{Aerosol}$ being independent of its uncertainty in the final period, but it is difficult to quantify the appropriate proportion. It is assumed here that 25% of the aerosol forcing total uncertainty variance is independent between the base and final periods, and we show the sensitivity to the alternate assumption of 50%.

Since the total 2011 uncertainty estimates in AR5 are for changes since 1750, it would generally be double counting to include a fixed element in uncertainty estimates (one that did not scale with the magnitude of the forcing) on account of forcing in 1750 or, by extension, in the base period used. However, exceptions are made for solar and volcanic forcings. Since solar forcing fluctuates about a mean, but is less accurately known in the past than in 2011, a fixed uncertainty element equal to the AR5 2011 uncertainty of ±0.05 Wm^{-2} is added thereto. For volcanic forcing, AR5 gives a symmetrical ±0.035 Wm^{-2} uncertainty range during 2008–2011. However, uncertainty in earlier periods is likely greater, with modest background stratospheric aerosol levels such as those over the last decade or so being undetectable. Therefore, a higher uncertainty of ±0.08 Wm^{-2} is used, composed of uncertainty of ±0.035 Wm^{-2} (per AR5) treated as fractional and ±0.072 Wm^{-2}, treated as fixed, added in quadrature. Separate random uncertainty realizations are applied in the base and final periods to sample the fixed uncertainty elements for each of $ERF_{Volcano}$ and ERF_{Solar}.

5.2 Computing surface temperature uncertainty distributions

We use the HadCRUT4 ensemble data for GMST, which provides 100 temperature realizations sampling uncertainties from systematic biases that preserve the correlation structure between time periods. The median for ΔT is calculated as the difference between the means over the final and base periods of the 100 HadCRUT4 ensemble members for the relevant years, with uncertainty in each period derived from the covariance matrix of observational uncertainty that is estimated using the 100 realizations together with HadCRUT4 estimates of types of uncertainty not sampled by the ensemble. The uncertainty estimates for GMST in the base and final periods, and internal variability in ΔT derived as set out in Section 3.3, and are all added in quadrature to give the uncertainty distribution for ΔT.

5.3 Computing heat uptake uncertainty distributions

Total heat uptake rates in the final and base periods and their associated uncertainties are estimated as set out in Section 3.2 and ΔQ is calculated as the difference in best estimates of the heat uptake between the final and base period. The distribution for total uncertainty in ΔQ is derived by adding in quadrature uncertainty estimates for the base and final periods together with, for both periods, the internal variability estimate detailed in Section 3.2.

5.4 Sampling from the uncertainty distributions

Two million samples are drawn randomly from each of the relevant normal or shifted-lognormal distributions that reflect the appropriate medians and uncertainties. For ΔT, where a single Gaussian uncertainty distribution has been derived for each base period – final period pair, the samples are drawn from the standard normal distribution and used to scale the median and standard deviation of the ΔT uncertainty distribution for each pair of periods. The same method is used for ΔQ. In the case of ΔF, the uncertainty distributions relating to the 2011 values for each of ERF_{WMGG}, $ERF_{Aerosol}$, $ERF_{BC\ on\ snow}$, $ERF_{Contrails}$, $ERF_{nonGABC}$, $ERF_{Volcano}$ and ERF_{Solar} are sampled. For those forcing components with uncertainty elements that are fixed and/or fractional but independent between the base and final periods,

those elements are sampled separately from the common fractional uncertainty, with different samples being drawn for the base period and for the final period.

5.5 Deriving ΔF and $F_{2\times CO2}$ samples from the sampled uncertainties of the forcing components

The resulting seven sets of ERF component samples incorporating common fractional uncertainty, and the sets representing fractional uncertainty elements that are independent between the base and final periods, are scaled to match the corresponding AR5-derived best estimate time series, averaged over each of the final periods and summed. The sets of samples representing fixed uncertainty elements are also added, without the scaling and averaging, thus producing samples of total ERF for each final period. In relation to fixed and to independent fractional uncertainty elements, the samples that were drawn for the final period are used. The process is repeated for the base periods, using the same set of samples as for the final periods in respect of the common fractional uncertainty components but using the separate samples that were drawn for the base period in respect of the fixed and the independent fractional uncertainty elements. Sample realizations of ΔF for each base period – final period pair are derived by deducting the sum of the sampled ERF components averaged over the relevant base period from the corresponding sum for the relevant final period.

For $ERF_{Aerosol}$, where the uncertainty distribution is a shifted lognormal, breaking up the total uncertainty (assumed to be entirely fractional) into two elements (common and independent) and then summing samples from them alters the distribution very slightly. A slight alteration to the $ERF_{Aerosol}$ 2011 median, leaving the 5– 95% range unchanged, is made so that the sum of the common and independent fractional components closely replicates the AR5 2011 $ERF_{Aerosol}$ median, 17–83% and 5–95% ranges given by the formulae in AR5 8.SM.7.

The same samples from the standard normal distribution that were scaled to sample the uncertainty distribution of ERF_{WMGG} are also used to scale $F_{2\times CO2}$ about its central estimate in AR5 of 3.71 Wm^{-2} in the same proportion, producing samples of $F_{2\times CO2}$ with matching uncertainty realizations to those for the ERF_{WMGG} samples.

5.6 Deriving best estimates and uncertainty ranges for ECS and TCR

For each base period – final period pair, estimates for ECS and TCR are calculated for each of the derived random realizations of ΔT, ΔF, ΔQ and $F_{2\times CO2}$. Sample histograms are computed using 0.01 K wide bins, and their distributions used to compute best estimates (medians) and uncertainty ranges for ECS and TCR. Any negative ECS or TCR sample estimates arising from the denominator in the calculation being negative are treated as infinitely high.

6. Results

ECS and TCR estimates based on each of the four preferred base period – final period combinations are presented in Table 4. The four estimates agree fairly closely for both ECS and TCR. The relevant results from Otto et al (2013) are shown for comparison.

Table 5 presents a sensitivity analysis. Sensitivity of ECS/TCR estimates to selection of base period is modest, provided allowance is made for volcanic forcing having a sub-unity efficacy, taken here as 0.55. The ECS best estimate using the 1995-2011 final period is then almost unchanged whether 1859-1882 or 1850-1900 is used as the base period; with unit volcanic efficacy the estimate falls from 1.64 K to 1.51 K when 1850-1900 is used. When using a volcanic efficacy of 0.55, the ranges of ECS and TCR best estimates from all nine base period – final period combinations are respectively 1.55–1.85 K and 1.18–

Base period	Final period	ECS best estimate [K]	ECS 17-83% range [K]	ECS 5-95% range [K]	TCR best estimate [K]	TCR 17-83% range [K]	TCR 5-95% range [K]
1859–1882	1995–2011	**1.64**	**1.25–2.45**	**1.05–4.05**	**1.33**	**1.05–1.80**	**0.90–2.50**
1850–1900	1987–2011	1.67	1.25–2.60	1.00–4.75	1.31	1.00–1.80	0.85–2.55
1850–1900	1971–2011	1.56	1.10–2.60	0.90–5.40	1.22	0.90–1.80	0.75–2.70
1930–1950	1995–2011	1.72	1.15–3.15	0.90–9.45	1.33	0.95–2.00	0.80–3.30
Otto et al (2013) results for comparison							
1860–1879	*2000–2009*	*2.00*	*1.50–2.80*	*1.20–3.90*	*1.33*	*1.05–1.65*	*0.90–1.95*
1860–1879	*1970–2009*	*1.91*	*1.30–3.05*	*0.95–5.00*	*1.36*	*0.95–1.90*	*0.75–2.55*

Table 4 Best estimates and uncertainty ranges for ECS and TCR using the base and final periods indicated. The preferred results are shown in bold. Best estimates are medians (50% points). Ranges are stated to the nearest 0.05 K. The final two lines, in italics, show the comparable results from Otto et al (2013) for the periods highlighted in that paper and used in AR5

1.35 K. Excluding 1971–2011, which is less well-matched for multidecadal internal variability, the TCR best-estimate range is only 1.30–1.35 K.

Sensitivity to variation in ΔF and ΔQ is assessed by re-estimating ECS and TCR using 1859–1882 and 1995–2011 data with ERF_{LUC} recentred on zero, which corresponds to a 0.10 Wm^{-2} increase in ΔF or reduction in ΔQ. In terms of the AR5 estimates of uncertainty in recent heat uptake, such a change represent one standard deviation for both 1995–2011 and 1987–2011. The effect (not shown in Table 5) of making the entire 2011 $ERF_{Aerosol}$ distribution 0.12 Wm^{-2} less negative, to match the average from the six satellite-based studies used to inform the AR5 composite best estimate, which produces a ~0.10 Wm^{-2} increase in ΔF, is the same as the effect of recentring ERF_{LUC} on zero. AR5 forcing uncertainty ranges are wide, reflecting a conservative approach. The effects of scaling down by 50% each of the three largest uncertainty ranges – those for $ERF_{Aerosol}$, ERF_{WMGG} and $ERF_{nonGABC}$ – are illustrated. The effect of raising the proportion of $ERF_{Aerosol}$ uncertainty that is treated as independent between the base and final periods from 25% to 50% is also shown. The effect of similarly doubling the independent part for $ERF_{nonGABC}$ (not shown) is minor. Increasing estimated internal variability in ΔT from 0.08 K to 0.12 K standard deviation has, as shown, only a modest effect.

Although final periods with end dates considerably before 2011 provide less well constrained ECS and TCR estimates, it is worthwhile investigating to what extent the low increase in GMST in the 21st century affects ECS and TCR best estimates. Accordingly, we estimated ECS and TCR using final periods from 1987 and 1971 to each of 2000, 2001, 2002 and 2003. As volcanic forcing is much higher than when 2011 is used as the end date, 1850–1900 is used as the best-matching base period. With a 1987 start date, the resulting ECS best estimates vary between 1.58 K and 1.70 K; those for TCR vary between 1.35 K and 1.37 K. With a 1971 start date, ECS best estimates vary between 1.45 K and 1.53 K; those for TCR vary between 1.20 K and 1.22 K. With a volcanic efficacy of 0.55 assumed, the ECS and TCR best estimates are all slightly lower.

Extending the final year to 2012, and estimating changes from 2011 in radiative forcings and in energy accumulation where necessary, has a negligible effect on ECS and TCR estimates.

Variation from 1859–1882 base period, 1995–2011 final period, main results case	ECS best estimate [K]	ECS 5-95% range [K]	TCR best estimate [K]	TCR 5-95% range [K]
Base case – no variations	1.64	1.0–4.1	1.33	0.9–2.5
Volcanic forcing efficacy set to 0.55	1.63	1.0–4.0	1.33	0.9–2.5
Base period 1850–1900; volcanic efficacy 1.0	1.51	1.0–3.2	1.23	0.8–2.1
Base period 1850–1900; volcanic efficacy 0.55	1.62	1.0–3.7	1.30	0.9–2.3
2011 ERF_{LUC} recentred on zero	1.54	1.0–3.5	1.27	0.9–2.3
$ERF_{Aerosol}$ uncertainty range scaled down to 50%	1.64	1.1–2.7	1.34	1.0–1.9
ERF_{WMGG} uncertainty range scaled down to 50%	1.63	1.0–3.8	1.33	0.9–2.5
$ERF_{nonGABC}$ uncertainty range scaled down to 50%	1.63	1.0–4.0	1.33	0.9–2.5
$ERF_{Aerosol}$ uncertainty independence raised to 50%	1.63	1.0–4.5	1.33	0.9–2.7
Internal variability of ΔT estimate raised by 50%	1.64	1.0–4.1	1.34	0.8–2.6

Table 5 Sensitivity of best estimates and uncertainty ranges for ECS and TCR using the 1995–2011 final period and, save where otherwise stated, the 1859–1882 base period. Ranges are stated to the nearest 0.1 K

7. Discussion

The estimates using the 1859-1882 base period and 1995-2011 final period combination are preferred because they not only involve the highest ΔT and ΔF, but also have minimal volcanic activity and well matched multidecadal variability.

The near identity of ECS and TCR best estimates based on warming from the late nineteenth century to both 1987–2011 and 1995–2011, and from 1930–1950 to 1995–2011, is impressive. The lower ECS and TCR estimates with a final period of 1971–2011 very likely reflect a less good match in multidecadal internal variability, and point to the desirability of climate sensitivity studies using data spanning periods over which trends are little affected by such variability.

The preferred ECS and TCR estimates are consistent with those from other recent instrumental-observation studies based on warming over the instrumental period that (i) either form their own inverse estimates of aerosol forcing using data resolved at least hemispherically, or use an estimate thereof that is consistent with the AR5 best estimate and (ii) do not use a prior distribution that strongly pushes their ECS or TCR estimate towards substantially higher values than those found here. Of the post-AR4 instrumental ECS estimates shown in Figure 10.20(b) of AR5, Aldrin et al. (2012), Lewis (2013) and Otto et al. (2013) are the only three meeting those requirements. Ring et al. (2012), cited in AR5 but omitted from Figure 10.20(b) as it did not provide an uncertainty range, also does so, as does Skeie et al. (2014). ECS estimates from four of these five studies are closely consistent with findings here. The two Otto et al ECS best estimates are slightly higher than those given here, in the case of that based on 2000–2009 data largely because Otto et al used an ocean heat content dataset that gave a high estimate of heat uptake over 2000–2009, as discussed in Section 2, in combination with a very low estimate of heat uptake in the base period. The Otto et al ECS estimate based on 1970–2009 data appears to be higher than our estimate based on 1971–2011 data primarily because their estimate used a base period

that substantially mismatched the final period in terms of volcanic activity, resulting in a lower ΔF value. Despite Otto et al giving higher best estimates for ECS than those here, its 95% uncertainty bounds for ECS are lower, reflecting primarily the much lower spread of forcing exhibited by CMIP5 AOGCMs than is implied by the AR5 forcing uncertainty estimates. Of the five studies, only Otto et al. (2013) and Skeie et al. (2014) gave TCR estimates; both have medians within 0.05 K of the 1.33 K obtained here.

It has been claimed (Rogelj et al, 2014) that the results of recent studies pointing to ECS being lower than the IPCC AR4 range (2–4.5 K) are strongly influenced by the small increase in observed warming during the last decade, although that is factually incorrect for all four of the studies it cited (Schmittner et al 2011, Aldrin et al 2012, Lewis 2013 and Otto et al 2013). Results here using final periods ending in 2000-2003 instead of 2011 do not support such claims either for ECS or for TCR.

Uncertainty in aerosol ERF is the largest contributor to imprecision in estimating ECS and TCR. Uncertainties in heat uptake and in WMGG forcing are substantially less important. Progress in reducing aerosol forcing uncertainty is therefore key to narrowing observationally-constrained estimates of climate sensitivity. Without a reduction in aerosol ERF uncertainty, additional observational data and extended time series may not lead to a major reduction in ECS and TCR estimation uncertainty.

Acknowledgements

We thank Gregory Johnson for supplying the data underlying Box 3.1, Figure 1 of AR5 and Steven Mosher and two reviewers for helpful comments.

References

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 23:253-271

Armour KC, Roe GH (2011) Climate commitment in an uncertain world. Geophysical Research Letters 38:L01707.

Armour KC, Bitz CM, Roe GR (2013) Time-varying climate sensitivity from regional feedbacks. J Clim 26:4518–4534

Bindoff NL, Stott PA, AchutaRao KM, Allen M, Gillett N, Gutzler D, Hansingo K, Hegerl G, Hu Y, Jain S, Mokhov I, Overland J, Perlwitz J, Sebbari R, Zhang X (2014) Detection and Attribution of Climate Change: from Global to Regional 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. Cambridge University Press, Cambridge

Bond, TC et al (2013) Bounding the role of black carbon in the climate system: A scientific assessment. J Geophys Res Atmos 118:5380–5552, doi:101002/jgrd50171

Boucher O, Randall D, Artaxo P, Bretherton C, Feingold G, Forster P, Kerminen V, Kondo Y, Liao H, Lohmann U, Rasch P, Satheesh SK, Sherwood S, Stevens B, Zhang X (2014) Clouds and Aerosols: 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. Cambridge University Press, Cambridge

Carslaw KS, Lee LA, Reddington CL, Pringle KJ, Rap A, Forster PM, Mann GW, Spracklen DV, Woodhouse MT, Regayre LA, Pierce JR (2013) Large contribution of natural aerosols to uncertainty in indirect forcing. Nature 503:67–71

Church JA, et al (2011) Revisiting the Earth's sea-level and energy budgets from 1961 to 2008. Geophys Res Lett 38:L18601

DelSole T, Tippett MK, Shukla J (2011) A significant component of unforced multidecadal variability in the recent acceleration of global warming J Clim 24:909-926

Domingues, CM, Church JA, White NJ, Gleckler PJ, Wijffels SE, Barker PM, Dunn JR (2008) Improved estimates of upper-ocean warming and multi-decadal sea-level rise. Nature 453:1090–3

Enfield DB, Mestas-Nunez AM, Trimble PJ (2001) The Atlantic Multidecadal Oscillation and its relationship to rainfall and river flows in the continental US. Geophys Res Lett 28:2077-2080

Forster P, Ramaswamy V, Artaxo P, Berntsen T, Betts R, Fahey DW, Haywood J, Lean J, Lowe DC, Myhre G, Nganga J, Prinn R, Raga DC, Schulz M, Van Dorland R (2007) Changes in Atmospheric Constituents and in Radiative Forcing: In 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 University Press, Cambridge

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 118:1139–1150

Fuglestvedt, J Terge Berntsen, Gunnar Myhre, Kristin Rypdal and Ragnhild Skeie (2008) Climate forcing from the transport sectors. PNAS 105(2):454-458

Gordon C, Cooper C, Senior CA, Banks H, Gregory JM, Johns TC, Mitchell JFB, Wood RA (2000) The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. Clim Dyn 16:147-168

Gregory JM, Stouffer RJ, Raper SCB, Stott PA, Rayner NA (2002) An observationally based estimate of the climate sensitivity. J Clim 15:3117–3121

Gregory JM, Forster PM (2008) Transient climate response estimated from radiative forcing and observed temperature change. J. Geophys. Res. 113.

Gregory JM et al. (2013) Climate models without pre-industrial volcanic forcing underestimate historical ocean thermal expansion. Geophys Res Lett 40(8):1600–1604

Hansen J et al (2005) Efficacy of climate forcings. J Geophys Res, 110: D18104, doi:101029/2005JD005776

Hartmann DL, Klein Tank AMG, Rusticucci M, Alexander LV, Brönnimann S, Charabi Y, Dentener FJ, Dlugokencky EJ, Easterling DR, Kaplan A, Soden BJ, Thorne PW, Wild M, Zhai PM (2014) Observations: Atmosphere and Surface: 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. Cambridge University Press, Cambridge

Ishii M, Kimoto M (2009) Reevaluation of historical ocean heat content variations with time-varying XBT and MBT depth bias corrections. J Oceanogr 65:287-299

Levitus S et al (2012) World ocean heat content and thermosteric sea level change (0-2000m) 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 26:7414-7429

Loeb, NG et al (2012 Observed changes in top-of-the-atmosphere radiation and upper-ocean heating consistent within uncertainty. Nature Geoscience 5:110–113

Lyman JM et al (2010) Robust warming of the global upper ocean. Nature 465:334-337

Lyman, JM., Johnson GC (2014) Estimating Global Ocean Heat Content Changes in the Upper 1800 m since 1950 and the Influence of Climatology Choice. J. Climate, 27, 1945–1957.

Masters T (2014) Observational estimate of climate sensitivity from changes in the rate of ocean heat uptake and comparison to CMIP5 models. Clim Dyn 42:2173-2181 DOI 101007/s00382-013-1770-4

Meinshausen M, Meinshausen N, Hare W, Raper SCB, Frieler K, Knutti R, Frame DJ, Allen MR (2009) Greenhouse gas emission targets for limiting global warming to 2°C. Nature 458:1158-1162

Meinshausen M, Smith SJ, Calvin K, Daniel JS, Kainuma MLT, Lamarque J-F, Matsumoto K, Montzka SA, Raper SCB, Riahi K, Thomson A, Velders GJM, van Vuuren DPP (2011) The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. Climatic Change 109(1-2):213-241

Meraner K, Mauritsen T, Voigt A (2013) Robust increase in equilibrium climate sensitivity under global warming. Geophys Res Lett 40:1-5 doi:101002/2013GL058118

Morice C P, 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 doi:101029/2011JD017187

Myhre G, Shindell D, Bréon F, Collins W, Fuglestvedt J, Huang J, Koch D, Lamarque J, Lee D, Mendoza B, Nakajima T, Robock A, Stephens G, Takemura T, Zhang H (2014) 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. Cambridge University Press, Cambridge

Otto A, Otto FEL, Boucher O, Church J, Hegerl G, Forster PM, Gillett NP, Gregory J, Johnson GC, Knutti R, Lewis N, Lohmann U, Marotzke J, Myhre G, Shindell D, Stevens B, Allen MR (2013) Energy budget constraints on climate response. Nature Geosci 6:415–416

Ring MJ, Lindner D, Cross EF, Schlesinger ME (2012) Causes of the global warming observed since the 19th century. Atmos Clim Sci 2:401–415

Roe GH, Armour KC (2011) How sensitive is climate sensitivity? Geophysical Research Letters 38:L14708.

Rogelj J, Meinshausen M, Sedlácek J, Knutti R (2014) Implications of potentially lower climate sensitivity on climate projections and policy. Environ Res Lett 9:doi:101088/1748-9326/9/3/031003

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. Geophysical Research Letters 41:1071--1078.

Schmittner A, Urban NM, Shakun JD, Mahowald NM, Clark PU, Bartlein PJ, Mix AC, Rosel-Mele A (2011) Climate sensitivity estimated from temperature reconstructions of the last glacial maximum. Science 334:1385–8

Schwartz, SE (2012) Determination of Earth's transient and equilibrium climate sensitivities from observations over the twentieth century: strong dependence on assumed forcing. Surveys in geophysics 33.3-4: 745-777.

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 Dynam 5:139–175

Smith DM, Murphy JM (2007) An objective ocean temperature and salinity analysis using covariances from a global climate model. J Geophys Res 112(C02022): DOI: 10.1029/2005JC003172

Smith S J, van Aardenne J, Klimont Z, Andres RJ, Volke A, Delgado Arias S (2011) Anthropogenic sulfur dioxide emissions: 1850–2005. Atmos Chem Phys 11:1101–1116

Stevenson DS, et al (2013) Tropospheric ozone changes, radiative forcing and attribution to emissions in the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP) Atmos Chem Phys 13:3063-3085

Tomassini L, Reichert P, Knutti R, Stocker TF, Borsuk ME (2007) Robust Bayesian uncertainty analysis of climate system properties using Markov chain Monte Carlo methods. J Clim 20:1239-1254

Tung K, Zhou J (2013) Using data to attribute episodes of warming and cooling instrumental records. PNAS 110:2058–2063

Willis, JK, Lyman JM, Johnson GC, Gilson J (2009) In Situ Data Biases and Recent Ocean Heat Content Variability. J. Atmos. Oceanic Technol., 26, 846–852

Wolter K, Timlin MS (1993) Monitoring ENSO in COADS with a seasonally adjusted principal component index. Proc of the 17th Climate Diagnostics Workshop, Norman, OK, NOAA/NMC/CAC, NSSL, Oklahoma Clim Survey, CIMMS and the School of Meteor, Univ of Oklahoma, 52-57

Wolter K, Timlin MS (2011) El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEIext). Intl J Climatol 31:1074–1087

Zelinka MD, Hartmann DL (2012) Climate Feedbacks and Their Implications for Poleward Energy Flux changes in a warming climate. J Clim 25:608–624

Zelinka, MD, Andrews T, Forster PM and Taylor KE (2014) Quantifying components of aerosol-cloud-radiation interactions in climate models, J. Geophys. Res. Atmos., 119, *doi:*10.1002/2014JD021710.