Why is estimating climate sensitivity so problematical?

Guest blog Nic Lewis

Introduction

Climate sensitivity estimates exhibit little consistency. As shown in the Introduction, Figure 1 of Box 12.2 of AR5[i] (reproduced here as Figure 1) reveals that 5–95% uncertainty ranges estimated for equilibrium climate sensitivity (ECS) vary from 0.6–1.0°C at one extreme (Lindzen & Choi, 2011), to 2.2–9.2°C at the other (Knutti, 2002), with medians[ii] ranging from 0.7°C to 5.0°C.

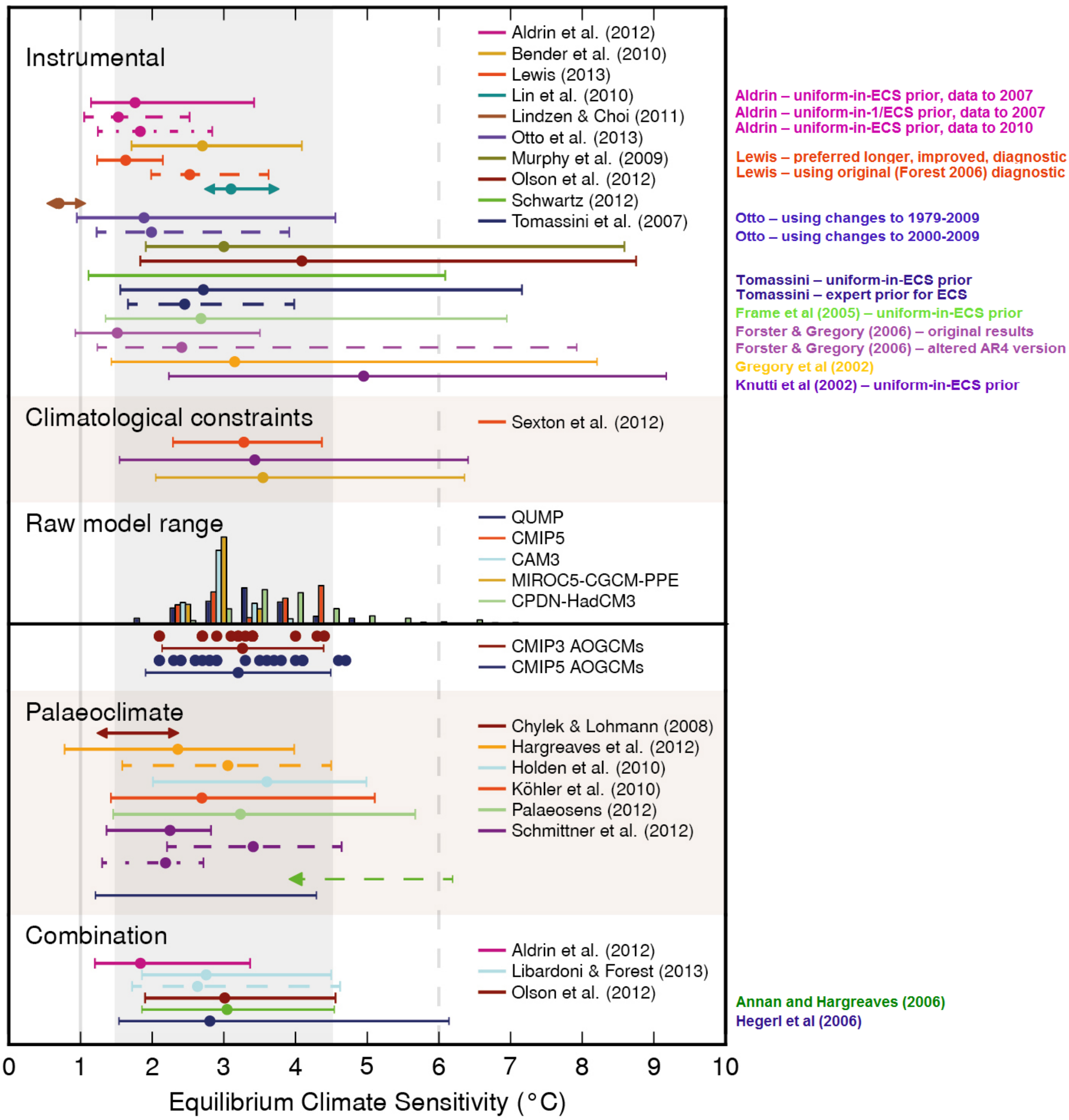

Figure 1. Annotated reproduction of Box 12.2, Figure 1 from AR5 WG1: ECS estimates
Bars show 5–95% uncertainty ranges for ECS, with best estimates (medians) marked by dots. Actual ECS values are given for CMIP3 and CMIP5 GCMs. Unlabelled ranges relate to studies cited in AR4.

The ECS values of CMIP5 general circulation or global climate models (GCMs) – as indicated by the dark blue dots in figure 1 – cover a narrower, but still wide range, of 2.1–4.7°C. So how should one weight the different lines of evidence, and the studies within them?

Climatological constraint studies

All climatological constraints ECS estimates cited in AR5 come from studies based on simulations by multiple variants of the UK HadCM3/SM3 GCM, the parameters of which have been systematically varied to perturb the model physics and hence its ECS values. These are called Perturbed Physics Ensemble (PPE) studies. Unfortunately, the HadCM3/SM3 model, maybe in common with other models, has a structural link – probably via clouds – between ECS and aerosol radiative forcing. As a result, at parameter settings that produce even moderately low ECS values, aerosol cooling is so high that the model climate becomes inconsistent with observations. See Box 1 in this document for details. Therefore, the AR5 climatological constraint studies cannot provide scientifically valid observationally-based ECS estimates: they primarily reflect the characteristics of the HadCM3 GCM.

Categories of study that AR5 downplays

AR5 considers all observational ECS estimates. It concludes, in the final paragraph of section 12.5.3 that estimates based on

- paleoclimate data reflecting past climate states very different from today
- climate response to volcanic eruptions, solar changes and other non-greenhouse gas forcings
- timescales different from those relevant for climate stabilization, such as the climate response to volcanic eruptions

are unreliable, that is, may differ from the climate sensitivity measuring the climate feedbacks of the Earth system today. Another example of estimates based on different timescales (in practice, short-term changes) is satellite measured variations in top-of-atmosphere (TOA) radiation. The discussion of that approach in section 10.8.2.2 refers to uncertainties in estimates of the feedback parameter and the ECS from short-term variations in the satellite period precluding strong constraints on ECS. AR5 also concludes in the final sentence of section 10.8.2.4 that paleoclimate estimates support only a wide 10–90% range for ECS of 1.0–6°C. I agree with these conclusions, certainly for current studies.

Instrumental studies based on multidecadal warming

In essence, the only observational estimates remaining are those based on instrumental observations of warming over multi-decadal periods. In the last two or three decades the anthropogenic signal has risen clear of the noise arising from internal variability and measurement/forcing estimation uncertainty. These studies are therefore able to provide narrower ranges than those from paleoclimate studies. A key change between the 2007 AR4 report and AR5 has been a significant reduction in the best estimate of aerosol forcing, which – other things being equal – points to a reduction in estimates of ECS. However, uncertainties remain large, with the aerosol forcing uncertainty being by some way the most important for ECS estimation.

Useful surface temperature records extend back approximately 150 years (the 'instrumental period'). Global warming 'in the pipeline', representing the difference between transient climate response (TCR), a measure of sensitivity over 70 years, and ECS, is predominantly reflected in ocean heat uptake, calculated from changes in sub-surface temperatures, records of which extend back only some 50 years.

In effect, estimates based on instrumental period warming compare measured changes in temperatures with estimates of the radiative forcing from greenhouse gases, aerosols and other agents driving climate change. Some do so directly through mathematical relationships, but most use relatively simple climate models to simulate temperatures, which can then be compared with observations as the model's parameters (control knobs) are varied. The idea is that the most likely values for ECS (and any other key climate system properties being estimated) are those that correspond to the model parameter settings at which simulations best match observations.

Whichever method is employed, GCMs or similar models have to be used to help estimate most radiative forcings and their efficacy, the characteristics of internal climate variability and maybe other ancillary items. But these uses do not rely on the ECS values of the models involved: GCMs with very different ECS values can provide similar estimates of effective forcings, internal variability, etc.[iii] However, some ECS and TCR studies were based on GCM-derived estimates of anthropogenic warming or recent ocean heat uptake rather than observations. Although those estimates may have taken observational data into account, it is unlikely that they fully did so.

I will consider studies in the Combination category in Figure 1, Box 12.2 together with those in the Instrumental category, since the combination estimates all include an instrumental estimate. I include the unlabelled AR4 Instrumental studies Frame et al (2005), Gregory et al (2002) and Knutti et al (2002) and the unlabelled AR4 Combination study Hegerl et al (2006). I exclude the Lindzen & Choi (2011) and Murphy et al (2009) studies, and also the unlabelled AR4 Forster & Gregory (2006) study, as they are based on satellite measured short-term variations in TOA radiation, an approach deprecated by AR5. (Two of these three studies actually give low, well-constrained ECS estimates.) I exclude Bender et al (2010) and the unlabelled AR4 Combination study Annan & Hargreaves (2006) since they involve the response to volcanic eruptions, an approach deprecated in AR5.

That leaves all the AR4 and AR5 Instrumental and Combination studies that involving estimating ECS from multidecadal warming. They are a mixed bag: AR5 includes sensitivity estimates from flawed observational studies that used unsuitable data, were poorly designed and/or employed inappropriate statistical methodology. Before considering individual studies, I will highlight two particular issues that each affect a substantial number of the instrumental-period warming studies.

Aerosol forcing estimation

Many of the observational instrumental-period warming ECS estimates that were featured in Figure 1, Box 12.2 of AR5, or TCR estimates featured in Figure 10.20.a of AR5, used values for aerosol forcing that either:

a) were consistent with the *AR4* estimate; this was substantially higher than the estimate, based on better scientific understanding and observational data, given in AR5;
b) reflected aerosol forcing levels in particular GCMs that were substantially higher than the best estimates given in AR5; or
c) were estimated along with ECS using *global* mean temperature data.

Any of these approaches will lead to an unacceptable, biased ECS (or TCR) estimation. This is obvious for a) and b). Regarding c), because the time-evolution of global aerosol forcing is almost identical to that from greenhouse gases, it is impossible to estimate both aerosol forcing – which largely affects the northern hemisphere – and ECS (or TCR) with any accuracy without separate temperature data for the northern and southern hemispheres.

On my analysis, ECS estimates from Olson et al (2012), Tomassini et al (2007) and the AR4 study Knutti et al (2002) are unsatisfactory due to problem c).

Inappropriate statistical methodology
Most of the observational instrumental-period warming based ECS estimates cited in AR5 use a 'Subjective Bayesian' statistical approach.[iv] The starting position of many of them – their prior – is that all climate sensitivities are, over a very wide range, equally likely. In Bayesian terminology, they start from a 'uniform prior' in ECS. All climate sensitivity estimates shown in the AR4 report were stated to be on a uniform-in-ECS prior basis. So are many cited in AR5.

Use of uniform-in-ECS priors biases estimates upwards, usually substantially. When, as is the case for ECS, the parameter involved has a substantially non-linear relationship with the observational data from which it is being estimated, a uniform prior generally prevents the estimate fairly reflecting the data. The largest effect of uniform priors is on the upper uncertainty bounds for ECS, which are greatly inflated.

Instead of uniform-in-ECS priors, some climate sensitivity estimates use 'expert priors'. These are mainly representations of pre-AR5 'consensus' views of climate sensitivity, which largely reflect estimates of ECS derived from GCMs. Studies using expert priors typically produce ECS estimates that primarily reflect the prior, with the observational data having limited influence.

ECS estimates from the majority of instrumental-period warming based studies – identified below – are seriously biased up by use of unsuitable priors, typically a uniform-in-ECS prior and/or an expert prior for ECS. Unusually, although Aldrin et al (2012) used a Subjective Bayesian method, because its ECS estimates are well constrained they are only modestly biased by the use of a uniform-in-ECS prior (although its estimate using a uniform-in-1/ECS prior appears to reflect the data better).

Which instrumental warming studies are unsatisfactory, and why?
I will give just very brief summaries of serious problems that affect named studies and render their ECS estimates unsatisfactory.

Frame (2005) – ocean heat uptake incorrectly calculated; uses GCM-estimated anthropogenic warming not directly observed temperatures; ECS estimate badly biased by use of a uniform prior for ocean effective diffusivity (a measure of heat uptake efficiency) as well as for ECS.

Gregory (2002) – external estimate of forcing increase used was under half the AR5 best estimate.

Hegerl (2006) – ECS estimate dominated by one derived from the Frame (2005) study.

Knutti (2002) – poor aerosol forcing estimation [see c) above]; used a very weak pass/fail test to compare simulations with observations; estimate biased up by erroneous ocean heat content data and use of uniform prior for ECS.

Libardoni & Forest (2013) – ECS estimates largely reflect the expert prior used; surface temperature data badly used; and the relationships of its estimates using different datasets are unphysical.

Lin (2010) –forcing increase is too small (ignores strong volcanic forcing at start of simulation period) and assumed TOA imbalance excessive. Non-standard treatment of deep ocean heat uptake.

Olson (2012) – poor aerosol forcing estimation [see c) above]. Instrumental estimate using uniform prior for ECS almost unconstrained; Combination estimate dominated by the expert prior used.

Schwartz (2012) – The upper, 3.0–6.1°C, part of its ECS range derives from a poor quality regression using one of six alternative forcings datasets; the study concluded that dataset was inconsistent with the underlying energy balance model.

Tomassini (2007) – poor aerosol forcing estimation [see c) above]; ECS estimates badly biased by use of a uniform prior for ocean effective diffusivity and alternative uniform and expert priors for ECS.

For anyone who wants more details, I have made available, here, a fuller analysis of all the AR5 instrumental-period-warming based studies shown in Box 12.2, Figure 1 thereof, including Combination studies.

Which instrumental warming studies are satisfactory?
After setting aside all those instrumental-period-warming based studies where I find substantive faults, only three remain: Aldrin et al (2012), Lewis (2013) [solid line Box 12.1 Figure 1 range using improved diagnostic only] and Otto et al (2013). These all constrain ECS well, with best estimates of 1.5–2.0°C. Ring et al (2012), cited in AR5 but not shown in Box 12.1 Figure 1 as it provided no uncertainty ranges, also appears satisfactory. Its best estimates for ECS varied from 1.45°C to 2.0°C depending on the surface temperature dataset used.

Transient climate response estimation
Turning to TCR estimates cited in AR5, the story is similar. The ranges from AR5 studies are shown in Figure 2. As for ECS, I will give very brief summaries of serious problems that affect named studies and render their ECS estimates unsatisfactory.

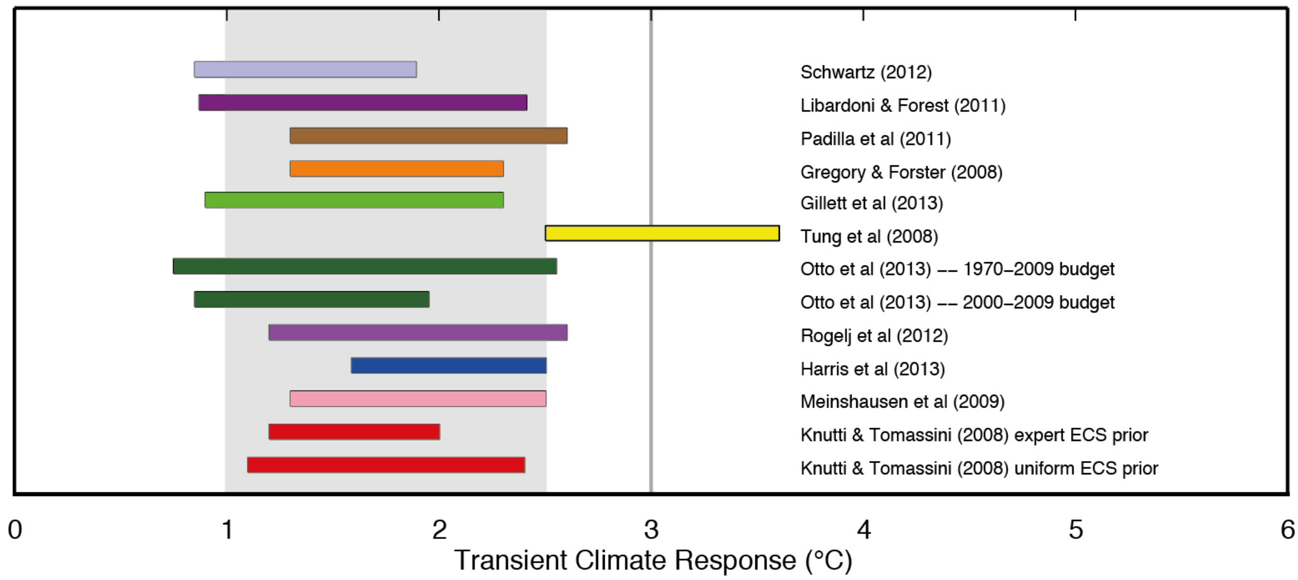

Figure 2. 5–95% TCR ranges from AR5 studies featured in Figure 10.20.a thereof

Libardoni & Forest (2011) – estimates largely reflect the ECS expert prior used; surface temperature data badly used; and the relationships of its estimates using different datasets are unphysical.

Padilla (2011) – poor aerosol forcing estimation [see c) above]; reducing uncertainty about aerosol forcing by using only post 1970 data lowers range from 1.3–2.6°C to 1.1–1.9°C. Its TCR estimate is sensitive to the forcing dataset and does vary logically with ocean mixed layer depth.

Gregory & Forster (2008) – regressed global temperature on anthropogenic forcing (excluding years with strong volcanism) over 1970–2006. That period coincided with the upswing half of the Atlantic Multidecadal Oscillation cycle, to which 0.1–0.2°C of the 0.5°C temperature rise was probably attributable. Regressing over 70 years using AR5 forcings gives a TCR best estimate of 1.3°C.

Tung (2008) – based on the response to the 11 year solar cycle. Section 10.8.1 of AR5 warns that its estimate may be affected by different mechanisms by which solar forcing affects climate.

Rogelj (2012) – neither a genuine observational estimate, nor published. The study imposes a PDF for ECS that reflects the AR4 likely range and best estimate, which together with the ocean heat uptake data used would have determined a PDF for TCR, with other data having very little influence.

Harris (2013) – an extension of the Sexton (2012) climatological constraint study to include recent climate change. Same problem: the study's TCR estimate mainly reflects the characteristics of the HadCM3/SM3 model, due to its structural link between ECS (& hence TCR) and aerosol forcing.

Meinshausen (2009) – TCR estimate is based on a PDF for ECS matching the AR4 best estimate and range. Finds a similar range using observations, but uses the high AR4 aerosol forcing estimate as a prior. The study appears to observationally constrain that prior weakly, probably because it attempts to constrain many more parameters than the 9 degrees of freedom it retains in the observations.

Knutti & Tomassini (2008) – uses same model setup, data and statistical method as the Tomassini (2007) ECS study, but estimates TCR instead. Same substantial problems as for that study.

I provide a more detailed analysis of AR5 TCR studies here.

On my analysis, only the Gillett et al (2013), Otto et al (2013) and Schwartz (2012) studies are satisfactory. Those studies give well-constrained best estimates for TCR of 1.3-1.45°C, averaging around 1.35°C.

Energy budget studies

It is instructive to consider the robust 'energy budget' method of estimating ECS (and by extension TCR), which involves fewer assumptions and less use of models than most others. In the energy budget method, external estimates – observationally based so far as practicable – of all components of forcing and heat uptake, as well as of global mean surface temperature, 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. Climate sensitivity may then be estimated as:

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

where F_{2xCO2} is the radiative forcing attributable to a doubling of atmospheric CO_2 concentration.

Strictly, Equation (1) provides an estimate of effective climate sensitivity rather than equilibrium climate sensitivity, according to the definitions in AR5. However, in practice the two terms are used virtually synonymously in AR5.

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). As AR5 states (p.920), Eq.(1) follows from conservation of energy. AR5 also points out that TCR represents a generic climate system property equalling the product of F_{2xCO2} (taken as 3.71 W/m^2 in AR5) 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 – as is the case over the instrumental period – then it likewise follows that:

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

The base and final periods each need to be at least a decade long, to reduce the effects of internal variability and measurement error. To obtain reliable and well-constrained estimation, one should choose base and final periods that capture most of the increase in forcing over the instrumental period and are similarly influenced by volcanic activity and internal variability, particularly multidecadal fluctuations. On doing so, best estimates for ECS and TCR using the forcing and heat uptake estimates given in AR5 and surface temperature records from the principal datasets are in line with those given above from studies that I do not find fault with. In fact, they lie in the lower halves of the 1.5–2.0°C ECS and 1.3-1.45°C TCR bands I quoted.

Note that Otto et al (2013), of which I am a co-author, was an energy budget study. It used average forcing estimates derived from CMIP5 GCMs rather than the AR5 estimates (which were not available at the time).

Raw model range

Before turning to evaluating the estimates of ECS from the CMIP3 (AR4) and CMIP5 (AR5) GCMs, I will first briefly discuss the TCR values that CMIP5 models exhibit. The AR5 projections of warming over the rest of this century should depend primarily reflect those TCR values.

Transient response is directly related to ECS, but lower on account of heat uptake by the climate system. CMIP5 GCMs have TCRs varying from 1.1°C to 2.6°C, averaging around 1.8°C – much higher than the sound observationally-based best estimates of 1.3-1.45°C. Moreover, about half the GCMs exhibit increases in transient sensitivity as forcing increases continue[v], so average CMIP5 projections of warming over the 21st century are noticeably higher than would be expected from their TCR values.

Feedbacks in GCMs

ECS in GCMs follows from the climate feedbacks they exhibit, which on balance amplify the warming effect of greenhouse gases.[vi] The main feedbacks in these models are water vapour, lapse rate, albedo and cloud feedbacks. Together, the first three of these imply an ECS of around 2°C. The excess of model ECS over 2°C comes primarily from positive cloud feedbacks and adjustments, with nonlinearities and/or climate state dependency also having a significant impact in some cases.

Problems with clouds

Reliable observational evidence for cloud feedback being positive rather than negative is lacking. AR5 (Section 7.2.5.7) discussed attempts to constrain cloud feedback from observable aspects of present-day clouds but concluded that "there is no evidence of a robust link between any of the noted observables and the global feedback".

Cloud characteristics are largely 'parameterised' in GCMs – calculated using semi-heuristic approximations rather than derived directly from basic physics. Key aspects of cloud feedback vary greatly between different models. GCMs have difficulty simulating clouds, let alone predicting how they will change in a warmer world, with different cloud types having diverse influences on the climate. Figure 3 shows how inaccurate CMIP5 models are in representing even average cloud extent; over much of the Earth's surface cloudiness is too low in most models.[vii]

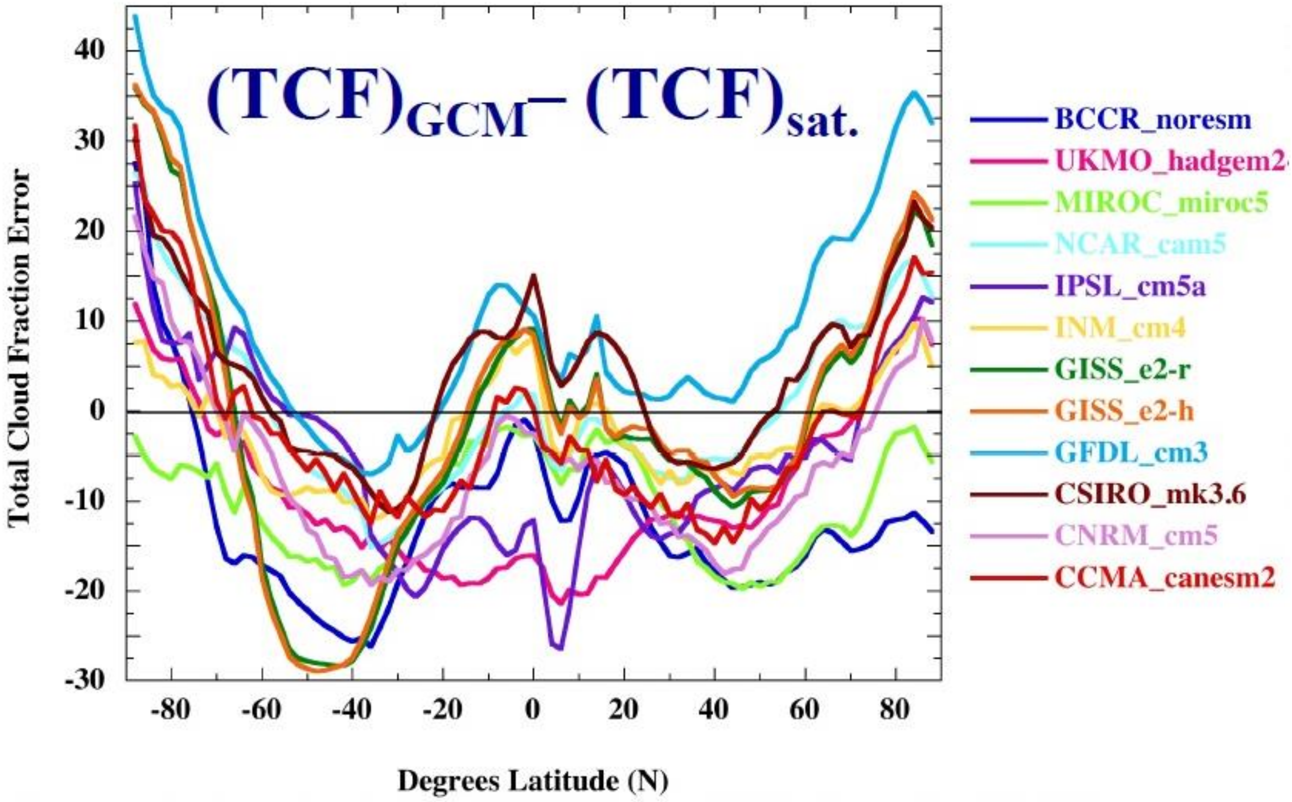

Figure 3. Error in total cloud fraction (TCF) for 12 CMIP5 GCMs. $(TCF)_{sat}$ = averaged MODIS and ISCCP2. Source: Patrick Frank, poster presentation at American Geophysical Union, Fall Meeting 2013: Propagation of Error and The Reliability of Global Air Temperature Projections

Although the overall effects of cloud behaviour on cloud feedback and hence on climate sensitivity are impossible to work out from basic physics and not currently well constrained by observations, the realism of GCM climate sensitivities can be judged from how their simulated temperatures have responded to the increasing forcing over the instrumental period. However, there is a complication.

Problems with aerosols

On average, GCMs exhibit significantly stronger (negative) aerosol forcing than the AR5 best estimate of -0.9 W/m^2 in 2011 relative to 1750. Averaged over CMIP5 models for which aerosol forcing has been diagnosed, its change over 1850 to 2000 appears to be around 0.4–0.5 W/m^2 more negative than per AR5's best estimate.[viii] In GCMs, much more of the positive greenhouse gas forcing would have been offset by negative aerosol forcing than per the AR5 best estimates, leaving a relatively weak average increase in net forcing. That depresses the simulated temperature rise over the instrumental period. With a weak forcing increase, GCMs needed to have high sensitivities in order to match the warming experienced from the late 1970s until the early 2000s. If aerosol forcing is actually smaller and the models had correctly reflected that fact, they would – given their high sensitivities – have simulated excessive warming.

If aerosol forcing is close to AR5's best estimate, there is little doubt that most of the models are excessively sensitive. But what if AR5's best estimate of aerosol forcing is insufficiently negative? The uncertainty range of the AR5 aerosol forcing estimate is very wide, and probably encompasses all GCM aerosol forcing levels. At present, one cannot say for certain that average GCM aerosol forcing is excessive.

Too fast warming once aerosol forcing stabilised

Fortunately, there is general agreement that aerosol forcing has changed little – probably by no more than ±0.15 W/m^2 – since the end of the 1970s. By comparison, over 1979–2012 other forcings increased by about 1.3 W/m^2. So by comparing model-simulated global warming since 1979 with actual warming, we can test whether the CMIP5 GCMs sensitivity is realistic without worrying too much about aerosol forcing uncertainty. Figure 4 shows that warming comparison over the 35 years 1979–2013, a period that is long enough to be used to judge the models. Virtually all model climates warmed much faster than the real climate, by 50% too much on average. Moreover, this was a period in which the main source of multidecadal internal variability in global temperature, the Atlantic Multidecadal Oscillation (AMO), had a positive influence (see, e.g., Tung and Zhou, 2012). Without its positive influence on the real climate, which was not generally included in GCM simulations, the average excess of CMIP5 model warming over actual would have been far more than 50%.

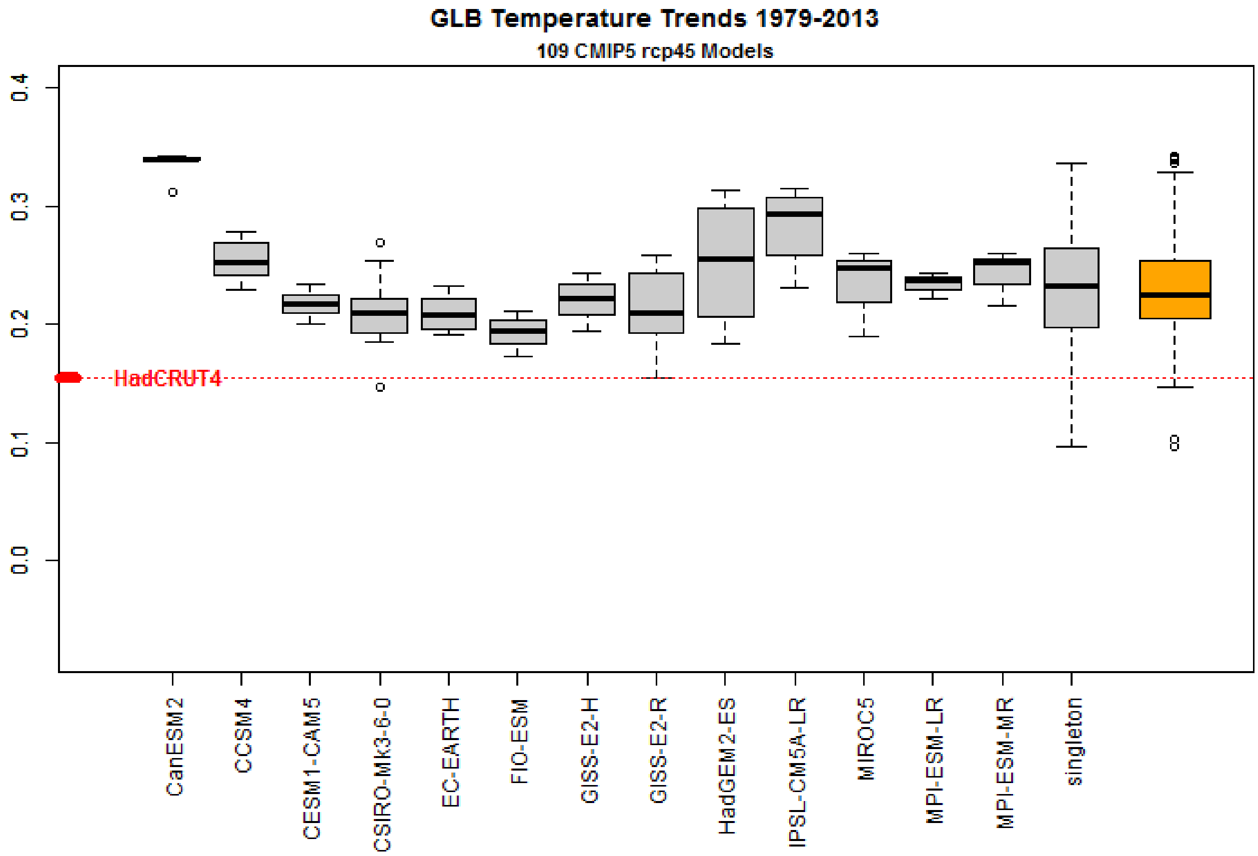

Figure 4. Modelled versus observed decadal global surface temperature trend 1979–2013 Temperature trends in °C/decade. Source: http://climateaudit.org/2013/09/24/two-minutes-to-midnight/. Models with multiple runs have separate boxplots; models with single runs are grouped together in the boxplot marked 'singleton'. The orange boxplot at the right combines all model runs together. The red dotted line shows the actual increase in global surface temperature over the same period per the HadCRUT4 observational dataset.

Over the slightly shorter 1988–2012 period, Figure 9.9 of AR5, reproduced here as Figure 5, shows an even more striking difference in trends in tropical lower tropospheric temperature over the oceans. The median model temperature trend (shown along the *x*-axis: the *y*-axis is not relevant here) is three times that of the average of the two observational datasets, UAH and RSS.

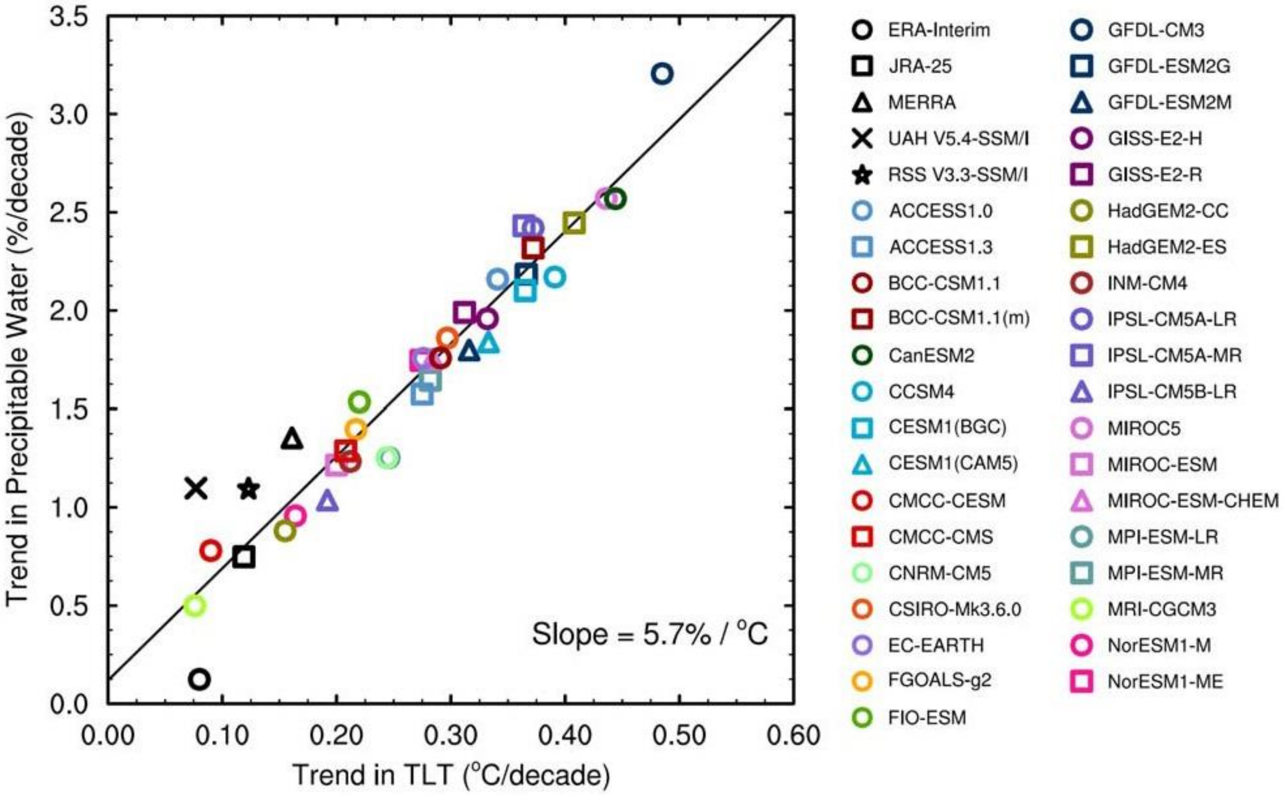

Figure 5. Reproduction of Figure 9.9 from AR5 WG1
Decadal trends for the 1988–2012 period in tropical (20°S to 20°N) lower tropospheric temperature (TLT) over the oceans are shown along the x-axis. Coloured symbols are from CMIP5 models. The black cross (UAH) and black star (RSS) show trends per satellite observations. Other black symbols are from model-based data reanalyses. All but two CMIP5 models exhibit higher TLT trends than UAH and RSS.

To summarise, the ECS and TCR values of CMIP5 models are not directly based on observational evidence and depend substantially on flawed simulations of clouds. Moreover, in the period since aerosol forcing stabilised ~35 years ago most models have warmed much too fast, indicating substantial oversensitivity. I therefore consider that little weight should be put on evidence from GCMs (and the related feedback analysis) as to the actual levels of ECS and TCR.

Conclusions
To conclude, I would summarise my answers to the questions posed in the Introduction as follows:

1. Observational evidence is preferable to that from models, as understanding of various important climate processes and the ability to model them properly is currently limited.

2. Little weight should be given to ECS evidence from the model range or climatological constraint studies. Of observational evidence, only that from warming over the instrumental period should be currently regarded as both reliable and able usefully to constrain ECS, in accordance with the conclusions of AR5. Studies that have serious defects should be discounted.

3. The major disagreement between ECS best estimates based on the energy budget, of no more than about 2°C, and the average ECS value of CMIP5 models of about 3°C, seems to me the main reason why the AR5

scientists felt unable to give a best estimate for ECS. All the projections of future climate change in AR5 are based on the CMIP5 models. Giving a best estimate materially below the CMIP5 model average could have destroyed the credibility of the Working Group 2 and 3 reports. As it is still difficult, given the uncertainties, to rule out ECS being as high as the CMIP5 average, I do not criticise the lack of a best estimate in AR5. However, I think a more forthright and detailed explanation of the reasons was called for. I would have liked a clear statement that most model sensitivities lay towards the top of the uncertainty range implied by the AR5 forcing and heat uptake estimates.

4. The soundest observational evidence seems to point to a best estimate for ECS of about 1.7°C, with a 'likely' (17-83%) range of circa 1.2–3.0°C.

5. Following a detailed analysis of all studies featured in AR5, the only TCR estimates that I consider significant weight should be given to are those from the Otto, Gillett and Schwartz studies.

6. The soundest observational evidence points to a 'likely' range for TCR of about 1.0–2.0°C, with a best estimate of circa 1.35°C.

Biosketch

Nic Lewis is an independent climate scientist. He studied mathematics and physics at Cambridge University, but until about five years ago worked in other fields. Since then he has been researching in climate science and in areas of statistics of relevance to climate science. Over the last few years he has concentrated mainly on the problem of estimating climate sensitivity and related key climate system properties. He has worked with prominent IPCC lead authors on a key paper in the area. He is also sole author of a recent paper that reassessed a climate sensitivity study featured in the IPCC AR4 report, showing that the subjective statistical method it used greatly overstated the risk of climate sensitivity being very high. Both papers are cited and discussed in the IPCC's recently released Fifth Assessment Report.

References

Aldrin, M., M. Holden, P. Guttorp, R.B. Skeie, G. Myhre, and T.K. Berntsen, 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.

Annan, J.D. and J.C. Hargreaves, 2006. Using multiple observationally-based constraints to estimate climate sensitivity. *Geophys. Res. Lett.,* 33: L06704.

Forest, C.E., P.H. Stone and A.P. Sokolov, 2006.Estimated PDFs of climate system properties including natural and anthropogenic forcings. *Geophys. Res. Lett*., 33: L01705

Forster, P.M.D., and J.M. Gregory, 2006. The climate sensitivity and its components diagnosed from Earth radiation budget data. *J.Clim.,*19: 39–52.

Frame D.J., B.B..B. Booth, J.A. Kettleborough, D.A. Stainforth, J.M. Gregory, M. Collins andM.R. Allen, 2005.Constraining climate forecasts: The role of prior assumptions*. Geophys. Res. Lett.*, 32, L09702

Gillett, N.P., V.K. Arora, D. Matthews, P.A. Stott, and M.R. Allen, 2013. Constraining the ratio of global warming to cumulative CO2 emissions using CMIP5 simulations. *J. Clim.*, doi:10.1175/JCLI-D-12–00476.1.

Gregory, J.M., R.J. Stouffer, S.C.B. Raper, P.A. Stott, and N.A. Rayner, 2002. An observationally based estimate of the climate sensitivity. *J. Clim.*,15: 3117–3121.

Gregory J.M. and P.M.Forster, 2008. Transient climate response estimated from radiative forcing and observed temperature change. *J.Geophys. Res.*, 113, D23105.

Harris, G.R., D.M.H. Sexton, B.B.B. Booth, M. Collins, and J.M. Murphy, 2013. Probabilistic projections of transient climate change. *Clim. Dynam.*, doi:10.1007/s00382–012–1647-y.

Hegerl, G.C., T.J. Crowley, W.T. Hyde, and D.J. Frame, 2006. Climate sensitivity constrained by temperature reconstructions over the past seven centuries. *Nature*;440: 1029–1032.

Knutti, R., T.F. Stocker, F. Joos, and G.-K. Plattner, 2002. Constraints on radiative forcing and future climate change from observations and climate model ensembles. *Nature*, 416: 719–723.

Knutti, R. and G.C. Hegerl, 2008. The equilibrium sensitivity of the Earth's temperature to radiation changes. *Nature Geoscience*;1: 735–743.

Lewis, N., 2013. An objective Bayesian, improved approach for applying optimal fingerprint techniques to estimate climate sensitivity. *J. Clim.,*26: 7414–7429.

Libardoni, A.G. and C.E.Forest, 2011. Sensitivity of distributions of climate system properties to the surface temperature dataset. *Geophys. Res. Lett.*; 38, L22705.

Libardoni, A.G. and C.E.Forest, 2013. Correction to 'Sensitivity of distributions of climate system properties to the surface temperature dataset'. *Geophys. Res. Lett.*; doi:10.1002/grl.50480.

Lin, B., et al., 2010: Estimations of climate sensitivity based on top-of-atmosphere radiation imbalance. *Atmos. Chem. Phys.*, 10: 1923–1930.

Lindzen, R.S. and Y.S. Choi, 2011. On the observational determination of climate sensitivity and its implications. *Asia-Pacific J. Atmos. Sci.*;47: 377–390.

Meinshausen, Malte, Nicolai Meinshausen, William Hare, Sarah C. B. Raper, Katja Frieler, Reto Knutti, David J. Frame, Myles R. Allen, 2009: Greenhouse gas emission targets for limiting global warming to 2°C. *Nature*, doi: 10.1038/

Murphy, D.M., S. Solomon, R.W. Portmann, K.H. Rosenlof, P.M. Forster, and T. Wong, 2009. An observationally based energy balance for the Earth since 1950. *J. Geophys. Res. Atmos.,*114: D17107.

Olson, R., R. Sriver, M. Goes, N.M. Urban, H.D. Matthews, M. Haran, and K. Keller, 2012. A climate sensitivity estimate using Bayesian fusion of instrumental observations and an Earth System model. *J. Geophys. Res. Atmos.,*117: D04103.

Otto, A., F. E. L. Otto, O. Boucher, J. Church, G. Hegerl, P. M. Forster, N. P. Gillett, J. Gregory, G. C. Johnson, R. Knutti, N. Lewis, U. Lohmann, J. Marotzke, G. Myhre, D. Shindell, B Stevens and M. R. Allen, 2013: Energy budget constraints on climate response. *Nature Geoscience*, 6, 415–416.

Ring, M.J., D. Lindner, E.F. Cross, and M.E. Schlesinger, 2012. Causes of the global warming observed since the 19th century. *Atmos. Clim. Sci.*, 2: 401–415.

Rogelj, J., M. Meinshausen and R. Knutti, 2012. Global warming under old and new scenarios using IPCC climate sensitivity range estimates. *Nature Climate Change*, 2, 248–253

Schwartz, S.E., 2012. Determination of Earth's transient and equilibrium climate sensitivities from observations over the twentieth century: Strong dependence on assumed forcing. *Surv.Geophys*., 33: 745–777.

Sexton, D.M. H., J.M. Murphy, M. Collins, and M.J. Webb, 2012. Multivariate probabilistic projections using imperfect climate models part I: outline of methodology. *Clim. Dynam.*, 38: 2513–2542.

Shindell, D.T. et al, 2013. Radiative forcing in the ACCMIP historical and future climate simulations, *Atmos. Chem. Phys.*, 13, 2939-2974

de Szoeke, S.P. et al, 2012. Observations of Stratocumulus Clouds and Their Effect on the Eastern Pacific Surface Heat Budget along 20°S. *J. Clim*, **25**, 8542–8567.

Tomassini, L., P. Reichert, R. Knutti, T.F. Stocker, and M.E. Borsuk, 2007. Robust Bayesian uncertainty analysis of climate system properties using Markov chain Monte Carlo methods. *J. Clim.*, 20: 1239–1254.

Tomassini, L.et al, 2013. The respective roles of surface temperature driven feedbacks and tropospheric adjustment to CO_2 in CMIP5 transient climate simulations. *Clim. Dyn*, DOI 10.1007/s00382-013-1682-3.

Tung, K-K and J Zhou, 2013. Using data to attribute episodes of warming and cooling in instrumental records. *PNAS*, 110, 6, 2058–2063

---

[i] References to AR5 are to the Working Group 1 report of the IPCC Fifth Assessment, except where the context requires otherwise.

[ii] The 50% probability point, which the target of the estimate is considered equally likely to lie above or below. All the best estimates I quote are medians, unless otherwise stated.

[iii] For instance, one can compute instantaneous radiative forcing (RF) for GHG without a GCM, using line-by-line calculations. But in order to estimate effective radiative forcing (ERF) one needs a GCM to compute how the atmosphere reacts to the presence of the GHG and what effect that has on the TOA radiative balance. Whilst the ratio of the derived ERF to RF will not be totally independent of the GCM's ECS, as a first approximation it will be. And in fact the estimated ratio is close to unity for most forcing agents.

[iv] Aldrin et al (2012), Libardoni & Forest (2013), Olson et al (2012), Tomassini et al (2007) and, of the unlabelled AR4 studies, Annan & Hargreaves (2006), Frame et al (2005), Hegerl et al (2006), Knutti et al (2002) and (dashed bar only) Forster & Gregory (2006).

[v] Figure 1 of Tomassini et al (2013) shows that the global mean temperature increase in the second 70 years of the "1pctCO2" experiment exceeds that in the first 70 years by significantly more than is accounted for by emerging "warming in the pipeline" for 8 of the 14 models analysed. Gregory and Forster (2008), Table 1 also showed a similar behaviour for between 5 and 10 (rounding of the stated ratios precludes precise enumeration) of the 12 models analysed.

[vi] Broadly, ECS = F_{2xCO2}/ (3.2 - Sum of feedbacks), 3.2 representing the Planck response of increased radiation from a warmer Earth.

[vii] A peer reviewed study, Szoeke et al (2012) likewise found that simulations of the climate of the twentieth century by CMIP3 models had ~50% too few clouds in the area investigated (south-eastern tropical Pacific ocean), and thus far too little net cloud radiative cooling at the surface.

[viii] Shindell et al (2013) estimated the average change in total aerosol forcing from 1850 to 2000 for the CMIP5 models it analysed at -1.23 W/m²; the corresponding best estimate in AR5 is -0.74 W/m².