A SENSITIVE MATTER

HOW THE IPCC BURIED EVIDENCE SHOWING GOOD NEWS ABOUT GLOBAL WARMING

Nicholas Lewis and Marcel Crok

Foreword by Professor Judith Curry

The Global Warming Policy Foundation
GWPF Report 13 (Long Version)

GWPF REPORTS

THE GLOBAL WARMING POLICY FOUNDATION

BOARD OF TRUSTEES

ACADEMIC ADVISORY COUNCIL

A SENSITIVE MATTER

How the IPCC buried evidence showing good news about global warming

Nicholas Lewis and Marcel Crok

Foreword by Professor Judith Curry

ISBN: 978-0-9573880-1-8

Contents

Foreword

The sensitivity of our climate to increasing concentrations of carbon dioxide is at the heart of the scientific debate on anthropogenic climate change, and also the public debate on the appropriate policy response to increasing carbon dioxide in the atmosphere. Climate sensitivity and estimates of its uncertainty are key inputs into the economic models that drive cost-benefit analyses and estimates of the social cost of carbon.

The complexity and nuances of the issue of climate sensitivity to increasing carbon dioxide are not easily discerned from reading the Summary for Policy Makers of the assessment reports undertaken by the Intergovernmental Panel on Climate Change (IPCC). Further, the more detailed discussion of climate sensitivity in the text of the full Working Group I reports lacks context or an explanation that is easily understood by anyone not actively reading the published literature.

This report by Nic Lewis and Marcel Crok addresses this gap between the IPCC assessments and the primary scientific literature, providing an overview of the different methods for estimating climate sensitivity and a historical perspective on IPCC's assessments of climate sensitivity. The report also provides an independent assessment of the different methods for estimating climate sensitivity and a critique of the IPCC AR4 and AR5 assessments of climate sensitivity. It emphasizes the point that evidence for low climate sensitivity is piling up. I find this report to be a useful contribution to scientific debate on this topic, as well as an important contribution to the public dialogue and debate on the subject of climate change policy.

I agreed to review this report and write this Foreword since I hold both authors of this report in high regard. I have followed with interest Nic Lewis' emergence as an independent climate scientist and his success in publishing papers in major peer-reviewed journals on the topic of climate sensitivity, and I have endeavored to support and publicize his research. I have interacted with Marcel Crok over the years and appreciate his insightful analyses, most recently as a participant in climatedialogue.org.

The collaboration of these two authors in writing this report has resulted in a technically sound, well-organized and readily comprehensible report on the scientific issues surrounding climate sensitivity and the deliberations of the IPCC on this topic.

While writing this Foreword, I considered the very few options available for publishing a report such as this paper by Lewis and Crok. I am appreciative of the GWPF for publishing and publicizing this report. Public accountability of governmental and intergovernmental climate science and policy analysis is enhanced by independent assessments of their conclusions and arguments.

Judith Curry
Atlanta, GA, USA
February 2014

Judith Curry is Professor and Chair of the School of Earth and Atmospheric Sciences at the Georgia Institute of Technology. She is a fellow of the American Meteorological Society, the American Association for the Advancement of Science, and the American Geophysical Union.

About the authors

Nic Lewis

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.

Marcel Crok

Marcel Crok is a freelance science writer based in Amsterdam, The Netherlands. He used to work as an editor for *Natuurwetenschap & Techniek*, a Dutch popular science magazine (recently this magazine has become the Dutch edition of *New Scientist*). In 2005 he published a long article about the infamous hockey stick graph and the criticisms of Stephen McIntyre and Ross McKitrick. He published a Dutch book about global warming in 2010. After Climategate and the turmoil surrounding the AR4 IPCC report the Dutch government asked him to review the fifth IPCC assessment report as an expert reviewer. With the leading Dutch climate institutes KNMI and PBL, Crok also set up an international discussion platform, climatedialogue.org, which facilitates constructive dialogues between scientists with different views.

Executive summary

1. The scientific part (WGI) of the fifth IPCC assessment report (AR5), published in final form in January 2014, contains some really encouraging information.[1] The best observational evidence indicates our climate is considerably less sensitive to greenhouse gases than climate scientists had previously thought. The clues and the relevant scientific papers are all mentioned in the full IPCC report. However, this important conclusion is not drawn in the full report – it is only mentioned as a possibility – and is ignored in the Summary for Policymakers (SPM).

2. Until AR5, for 30 years the scientific establishment's best estimate and their uncertainty range for climate sensitivity had hardly changed. The best estimate for equilibrium climate sensitivity (ECS) started and ended at 3°C and the uncertainty range[2] generally had a lower bound of 1.5°C and an upper bound of 4.5°C.[3] However, several recent studies give best estimates of between 1.5°C and 2°C, substantially lower than most earlier studies indicated.

3. In the IPCC Fourth Assessment Report (AR4), the empirical estimates of climate sensitivity were largely based not only on data that has now been superseded, but also on an inappropriate statistical basis that biased them towards higher values, thus making the global warming problem appear 'worse'. In AR5, many studies still use inappropriate data and/or statistical methodology. However, there is now a body of empirical estimates of climate sensitivity, prepared using sound methodology and appropriate data, that give substantially lower values – both of long-term warming and of transient warming towards the end of this century – than climate model simulations.

4. Since the last IPCC report was prepared greenhouse gas concentrations have continued to increase, yet global temperatures have not risen; more importantly, estimates of the cooling efficacy of aerosol pollution have been cut. This combination of factors is indicative of the climate system being less sensitive to greenhouse gases than previously appeared to be the case. But the new evidence about aerosol cooling is not reflected in the computer climate models.

---

[1] The accepted final draft of the AR5 Working Group I report and the approved version of the Summary for Policymakers (SPM) were published in September 2013. Corrected final versions of the SPM and the full AR5 WGI report were released in January 2014.

[2] 'Likely', defined as the central two-thirds probability in the last two IPCC reports; until then it was not defined probabilistically.

[3] The fourth IPCC assessment report, published in 2007, increased the lower bound to 2°C.

5. Global climate models used to predict future climate change still generate model climate sensitivities in the range 2–4.5°C, averaging just over 3°C. Large parts of the IPCC reports are built around the computer model simulations. Almost all the projections of future climate change are based on them,[4] and a complete chapter is devoted to model performance. Admitting in the IPCC report that the best observationally-based estimates[5] of climate sensitivity are now only 1.5–2°C would imply that large parts of the AR5 report are out of line with the latest scientific evidence.

6. In our view, the IPCC WGI scientists were saddled with a dilemma. How should they deal with the discrepancy between climate sensitivity estimates based on models and sound observational estimates that are consistent with the new evidence about aerosol cooling? In conjunction with governments – who have the last say on the wording of the SPM – they appear to have decided to resolve this dilemma in the following way. First, they changed the 'likely' range for climate sensitivity slightly. It was 2–4.5°C in AR4 in 2007. They have now reduced the lower bound to 1.5°C, making the range 1.5–4.5°C. By doing this they went some way to reflect the new, lower estimates that have been published recently in the literature.

7. They also decided not to give a best estimate for climate sensitivity. The tradition of giving a best estimate for climate sensitivity goes all the way back to the Charney report in 1979, and all subsequent IPCC reports (except the third assessment report in 2001) gave one as well. In AR4 the best estimate was 3°C. At the time of approval of the SPM by governments in September 2013, the decision not to give a best estimate for climate sensitivity was mentioned only in a footnote in the SPM, citing 'a lack of agreement on values across assessed lines of evidence and studies'. Only in the final report, published in January 2014, was a paragraph added in the Technical Summary giving slightly more explanation.

8. At a minimum, the SPM should have given a more informative explanation of the decision to widen the ECS 'likely' range and not give any best estimate for ECS. That could have taken the form of a straightforward statement that the best-quality observational evidence, based on improved estimates of the effects of aerosol pollution and the extended record of warming now available,

---

[4] Projected warming increases less than proportionally with ECS due to the moderating effect of heat uptake by the ocean. Projected warming in the models could conceivably be in line with observational evidence despite their ECS not being so. But it is not.

[5] Observationally-based methods do involve some limited use of models, but the ways they are used to help derive climate sensitivity estimates from observations differ greatly from the way global climate models are used to produce sensitivity estimates.

points to a best estimate for ECS of 2°C or slightly less, while evidence from global climate models still suggests that it is about 3°C or slightly more. We – the authors of this report – were both expert reviewers of AR5 and in our review comments suggested that the IPCC should go further and give separate ranges for climate sensitivity based on models and on high quality observational studies.

9. In this report we suggest that the new observationally-based 'likely' range could be 1.25–3.0°C, with a best estimate of 1.75°C.[6] If the IPCC had made that change – which would have been in line with the best quality scientific evidence available – it would have been picked up by all the major news outlets in the world as one of the major, if not the major, outcomes of the report. And rightly so.

10. In AR5 the IPCC felt even more certain (95% certain, compared to 90% in AR4) that humans have caused most (more than 50%) of the warming since 1950. The media treated this as the major conclusion of AR5, but it is in fact a relatively trivial finding. The high-quality observationally-based estimates for climate sensitivity discussed in this report assume that virtually all the measured warming (not just since 1950, but over the last 100–150 years) is due to humans. The far more important question now is how much warming is likely in the future under various scenarios.

11. Transient climate response (TCR), a measure of warming from a doubling of carbon dioxide (CO_2) over a seventy-year period, reflects ocean heat uptake efficiency as well as climate sensitivity and is often seen as providing a better guide to warming over the twenty-first century than ECS.[7] AR5 lowers the 10–90% range for TCR of 1–3°C established in AR4 to a 'likely' range of 1–2.5°C. In this report, we suggest that an observationally-based 'likely' range for TCR could reasonably be 1–2°C, with a best estimate of 1.35°C. The average TCR for global climate models is much higher, at just under 2°C.

12. These lower, observationally-based estimates for climate sensitivity and TCR suggest that considerably less warming and sea level rise is to be expected in the future than the model projections imply. Projected future warming based on the best observationally-based estimate of TCR is 40–50% lower than

---

[6] This is based on giving precedence to high-quality estimates that use a long period of instrumental temperature data, in line with AR5's appraisal of the different types of estimate, and discounting studies with identified substantial failings.

[7] However, sea level response depends more on the relationship between ECS and TCR than on TCR itself.

the IPCC's model-based projected warming, and on the IPCC's second highest emissions scenario cumulative warming would still be around the international target of 2°C in 2081–2100.

13. Our criticisms are directed at the IPCC as an organisation,[8] on the constraints its process imposes, and on the excessive emphasis put on projections and other results derived from climate models. The scientists' hands were largely tied; the scopes and even titles of the various chapters had already been determined. Even discriminating between models would have been awkward politically.

14. The purpose of the IPCC is 'to assess on a comprehensive, objective, open and transparent basis the scientific, technical and socio-economic information relevant to understanding the scientific basis of risk of human-induced climate change'.[9] We believe that, due largely to the constraints the climate model-orientated IPCC process imposed, the WGI report and the SPM failed to reflect satisfactorily such an assessment in the case of climate sensitivity and TCR, arguably the most important parameters in the climate discussion.

---

[8] The IPCC is not a research organisation, but its assessment report process significantly influences research carried out by climate scientists, in particular that involving simulations by climate models.

[9] http://www.ipcc.ch/organization/organization_history.shtml.

Introduction

At the end of September2013 the UN Intergovernmental Panel on Climate Change (IPCC) launched the first and most important part of its fifth assessment (AR5): the Working Group I (WGI) report, entitled *Climate Change 2013, The Physical Science Basis*.[10] Leaks of drafts of the document had been publicised in the media for some time. The major conclusion of the report, presented in the Summary for Policymakers (SPM), was therefore no surprise to many. The IPCC felt even more certain about the role of humans on the climate than in its last report in 2007. The SPM claims that scientists are now 95% certain (up from 90% in 2007) that most of the warming since 1950 is due to human influences.[11]

This report is a reaction to the AR5 WGI report. We – the authors of this report – were both expert reviewers of AR5. Many of our comments were related to climate sensitivity, a key parameter in global warming discussions. This report focuses on how AR5 dealt with climate sensitivity.

Put very simply, if the climate is very sensitive to greenhouse gases and therefore climate sensitivity is high, then we can expect substantial warming in the coming century if greenhouse gas emissions are not severely reduced. If climate sensitivity is low, then future warming will be substantially lower, as will the rise in sea level.

Climate sensitivity is defined as the amount of global surface warming that occurs when the concentration of CO_2 in the atmosphere doubles. The term generally refers to the rise in temperature once the climate system has fully warmed up, a process taking over a thousand years due to the enormous heat capacity of the ocean. This so-called 'equilibrium climate sensitivity' (ECS), is the traditional and still most widely used measure. In practice what is more commonly estimated[12] is 'effective climate sensitivity', a close approximation

---

[10] The final WGI report, published on 30 January 2014, can be downloaded freely at http://www.climatechange2013.org/ . There were a fair number of changes, mostly minor, in the full report from the accepted version released at the end of September 2013. Minor corrections were also made to two sections of the SPM approved in September 2013. References to AR5 in this report should be read as referring to the AR5 WGI report except where the context requires otherwise.

[11] There is some confusion about this major conclusion. AR5 repeated the claim of AR4 that it is very likely (90% certain) that anthropogenic greenhouse gases are responsible for more than 50% of the warming since 1950. They increased the level to extremely likely (95% certain) though for the broader term 'human influences'. This includes other human influences such as soot, sulphate aerosols and land-use changes. According to AR5's best estimates, all warming since 1950 is due to anthropogenic influences.

[12] Including for global climate models. Note that by convention equilibrium climate sensitivity excludes adjustment by slow components of the climate system (e.g. ice sheets, vegetation).

to ECS that it is more practical to work with. The two terms are treated as synonymous in this report, as in effect they are in AR5.

A shorter-term measure of sensitivity, transient climate response (TCR), represents the extent of global warming over a 70 year timeframe during which CO_2 concentrations double.[13] TCR can be estimated more easily than ECS, and is more relevant to projections of warming – although not sea level rise – over the rest of this century.[14] We show estimates for both ECS and TCR later in this report.[15]

One could argue that the concept of climate sensitivity is rather simplistic. However, studies utilising complex global climate models indicate that the changes in many climatic variables of interest that are projected to occur at a particular increase in global mean temperature scale pro rata with different changes in global mean temperature.[16] That supports the usefulness of a global climate sensitivity measure.

Whilst a lot can be said about the relevance of the climate sensitivity concept, the fact is that it has played and still plays a key role in the debates about global warming, not only in scientific but also in political discussions.

In the international policy arena, the ultimate, two-decade-old goal is to limit global warming to a level that prevents 'dangerous human interference' with the climate, in the words of the UN Framework Convention on Climate Change. In recent years this has been defined – somewhat arbitrarily – as preventing warming to more than 2°C above preindustrial temperatures. We are already about 0.8°C of the way to this level of warming and have only 1.2°C to go. With a climate sensitivity of 3°C, consistent with climate models, 2°C of warming will very probably be reached later this century, depending mainly on how quickly emissions of greenhouse gases rise.

---

[13] The increase in CO_2 is specified to occur at a constant compound rate over the period, but modest fluctuations in the rate are unimportant. Estimation of TCR is unaffected by the actual rate of increase provided that the increase in global temperature is scaled appropriately, and TCR is little affected by moderate variations in the ramp period: between 60 and 80 years, at least.

[14] Although TCR is easier to estimate, unlike ECS it does not have a useful interpretation in terms of the physics of the climate system. TCR is lower than ECS because heat going into the ocean contributes to the value of ECS but not to TCR.

[15] The warming influence of CO_2 increases with the logarithm of its concentration. ECS and TCR can be used to work out the global surface temperature rise from a change in CO_2 concentration other than a doubling by scaling them pro rata to the change in $\log_2(CO_2$ concentration).

[16] Harris et al. (2006) found that the equilibrium spatial response pattern to a doubling of CO_2 concentration provided a good approximation to the pattern throughout a period of increasing CO_2 concentration. See also Section 12.4.2 of AR5 WGI.

The scientific validity of the two-degree target has been questioned.[17] For example Jaeger (2011) noted that:

> The 2° limit has emerged nearly by chance, and it has evolved in a somewhat contradictory fashion: policy makers have treated it as a scientific finding, scientists as a political issue. It has been presented as a threshold separating a domain of safety from one of catastrophe, and as an optimal strategy balancing costs and benefits. We propose to use it as a focal point in a coordination game, where a multitude of actors need to find a new coordination equilibrium in the face of climate risks.

Tol (2007) concluded that 'this target is supported by rather thin arguments, based on inadequate methods, sloppy reasoning, and selective citation from a very narrow set of studies'. Nevertheless, in a very apt 1998 paper Van der Sluijs noted that the concept of climate sensitivity 'acts as an "anchor" that fixes the scientific basis for the climate policy debate'.[18] For this reason the concept of climate sensitivity and derivative measures such as TCR remain very important, both in the scientific and policy arenas.

Unexpected decision in AR5

For over thirty years international assessments, including those of the IPCC, have presented both an uncertainty range and, generally, a best estimate for ECS. In most cases, the uncertainty range has been 1.5–4.5°C and the best estimate 3°C. In AR4[19] the range was adjusted slightly upwards to 2–4.5°C,[20] but AR5 reduced the lower bound down to 1.5°C, returning to the earlier range of 1.5–4.5°C for ECS and in effect admitting that the assessment in AR4 was suspect. However, AR5 gave no best estimate for ECS.

Given the importance of this decision one would have expected the SPM, and the accepted version of the full AR5 report, released a few days later, to have gone into some detail about the reasons for not giving a best estimate. However, this was not the case. The policymaker and interested reader are left with footnote 16 in the SPM which says:

---

[17] Jaeger and Jaeger (2011); Tol (2007).

[18] Van der Sluijs et al. (1998).

[19] The fourth IPCC assessment report, published in 2007.

[20] 'Likely' range defined as 66% or higher probability – implicitly a 17–83% probability range – in AR4 and AR5; until then not defined probabilistically.

> 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.

The value of ECS is arguably the most important parameter in climate science, and the decision not to offer any guidance as to whether its best estimate lies towards the bottom, in the middle or towards the top of the 'likely' range was unexpected.[21] To find only a limited explanation and then only in a footnote is rather surprising.

The full paragraph about climate sensitivity reads as follows (our emphasis):[22]

> The equilibrium climate sensitivity quantifies the response of the climate system to constant radiative forcing on multi-century time scales. It is defined as the change in global mean surface temperature at equilibrium that is caused by a doubling of the atmospheric CO_2 concentration. Equilibrium climate sensitivity is likely in the range 1.5°C to 4.5°C (high confidence), extremely unlikely less than 1°C (high confidence), and very unlikely greater than 6°C (medium confidence). The lower temperature limit of the assessed likely range is thus less than the 2°C in the AR4, but the upper limit is the same. *This assessment reflects improved understanding, the extended temperature record in the atmosphere and ocean, and new estimates of radiative forcing.* {TFE6.1, Figure 1; Box 12.2}

The clue behind the decision not to give a best estimate is contained in the final sentence of this paragraph, which we believe few policy makers will have noted. With this sentence the IPCC indicates that it had to reduce the lower bound for ECS to 1.5°C based on 'improved understanding, the extended temperature record in the atmosphere and ocean, and new estimates of radiative forcing'. We agree with the above sentence but we believe the consequences of the 'improved understanding' are further reaching than the AR5 report would lead us to believe. In this report we will explain why.

---

[21] The AR5 WGI Second Order Draft stated about ECS, in Box 12.2, 'The most likely value remains near 3°C.'

[22] In Section D.2 of the SPM. The concept of radiative forcing is explained in the 'Energy budget ECS estimates' subsection; see p. 17.

History of climate sensitivity estimates

The concept of climate sensitivity goes all the way back to the work of Arrhenius (1896), one of the founding fathers of the greenhouse theory, who first considered the effect of doubling CO_2 concentrations on the atmosphere. Later, in the 1960s and 1970s, computations with more comprehensive models confirmed the calculations of Arrhenius and the concept of climate sensitivity became firmly established.[23]

A National Academy of Sciences' report in 1979 (the Charney report) is regarded as the first major assessment of climate sensitivity. By that time, estimates of climate sensitivity were already based on numerical climate models, so-called general circulation models (GCMs; also known as global climate models). At the time of the Charney report only two models were available. One had a climate sensitivity of 2°C and the other of 4°C, explaining the best estimate of 3°C.

Table 1: Evolution of equilibrium climate sensitivity estimates in the last 35 years and the range for transient climate response since 2001

	ECS Range (°C)	**ECS Best estimate** (°C)	**TCR Range** (°C)
Charney Report 1979	1.5–4.5	3.0	
NAS Report 1983	1.5–4.5	3.0	
Villach Conference 1985	1.5–4.5	3.0	
IPCC First Assessment 1990	1.5–4.5	2.5	
IPCC Second Assessment 1995	1.5–4.5	2.5	
IPCC Third Assessment 2001	1.5–4.5	None given	1.1–3.1[a]
IPCC Fourth Assessment 2007	2.0–4.5[b]	3.0	1.0–3.0[c]
IPCC Fifth Assessment 2013	1.5–4.5[d]	None given	1.0–2.5[d]

[a] Range for AOGCMs; [b] Likely (17–83%) range; prior to AR4, ranges were not clearly defined in probabilistic terms. [c] 10–90% range. [d] Likely range.

Table 1 shows the evolution of both the range and the best estimate of ECS over the last 35 years. As one can see, not much has changed. The table also shows the evolution of the range for TCR since it was first given in 2001. Best estimates of TCR have not been published in any of these reports.

---

[23] Schlesinger et al. (2007).

GCM simulations have always played a key role in determining ECS. Originally GCM estimates of ECS did not fully model the ocean, but now GCMs with coupled atmosphere and ocean (AOGCMs) are used. AOGCMs have varying climate sensitivities but their average ECS is around 3°C, close to the value that has generally been the best estimate for the last 35 years. As one can see from the table, ECS best estimates have always fallen approximately in the middle of the range given.

The TCR of climate models can also be calculated from the model simulations. The average TCR of current AOGCMs is 1.8°C or so.

Van der Sluijs (1998) considered the reasons why the range for climate sensitivity has changed so little over a period in which the science has evolved enormously.[24] He concluded that the range was only partly determined by the science itself and that many other factors played a role. One of these was 'a need to create and maintain a robust scientific basis' for policy action.

As this report will make clear, Van der Sluijs's conclusions about the period prior to 1998 apply just as much to the present day. However, we will argue that the observational evidence supporting a substantial change in both the range and the best estimate for climate sensitivity is now so strong that any serious scientific assessment should discuss it.

Observations indicate a low climate sensitivity

Since the Charney report of 1979, GCMs have been an important tool for estimating climate sensitivity. Indeed, until the final years of the twentieth century the emerging anthropogenic signal was too small in relation to the noise of internal climate variability and measurement error for reliable direct observational estimation of ECS.

However, since then the signal has become stronger and it has become possible to derive well-constrained estimates of ECS using observational data from the instrumental period; that is, the period since around 1850, the date after which sufficient temperature measurements existed to combine them into a global average temperature. The IPCC calls such estimates 'Instrumental'.[25]

---

[24] Van der Sluijs et al. (1998).

[25] Observationally-based ECS estimation methods do involve some use of models, but the ways they are used to help derive climate sensitivity estimates from observations are very different from the way AOGCMs are used to produce sensitivity estimates.

Energy budget ECS estimates

In 2002, the UK scientist Jonathan Gregory and colleagues published a paper that set out a simple and straightforward method of deducing climate sensitivity from observations.[26] The authors described the advantage of their new method in their abstract:

> Because the method does not use the climate sensitivity simulated by a general circulation model, it provides an independent observationally-based constraint on this important parameter of the climate system.

Because of the importance of this robust 'energy budget' method of estimating climate sensitivity, we will explain it in some detail.

In equilibrium, incoming radiation from the sun at the top of the atmosphere (TOA) is balanced partly by reflected solar radiation but mainly by infrared radiation from the atmosphere and, to an extent, direct from the surface. When the concentration of greenhouse gases rises, making the atmosphere more opaque to infrared radiation, or when other drivers of global warming increase, a TOA radiation imbalance results. In other words, assuming surface temperature does not rise to compensate, less infrared radiation goes out than the net amount of sunlight that comes in. This imbalance is called the (radiative) forcing on the climate system (RF, measured in watts per square metre: W/m^2 or Wm^{-2}).

Suppose between two separate periods[27] one measures the changes in:

- mean forcing and
- the rate of increase in the Earth's climate system heat content.

Since during each period energy must be conserved, the difference between these two changes must have been counteracted by the increasing radiation caused by a rise in mean global surface temperature. From knowledge of all

---

[26] Gregory et al. (2002).

[27] Typically respectively early and late in the instrumental period, and each at least a decade long to minimise the uncertainty arising from the effects of natural internal climate system variability, which must be allowed for.

these changes, together with the forcing caused by a doubling of atmospheric CO_2 concentration ($F_{2\times \mathrm{CO}_2}$), one can deduce ECS.[28]

The energy budget method is also described in AR5, where it is pointed out that the calculation of ECS involved follows from the conservation of energy.[29] AR5 puts it well:

> ECS= $F_{2\times \mathrm{CO}_2}/\alpha$, where α is the sensitivity parameter representing the net increase in energy flux to space per degree of warming given all feedbacks operating on these timescales. Hence, by conservation of energy, ECS= $F_{2\times \mathrm{CO}_2} \times \Delta T/(\Delta F - \Delta Q)$, where ΔQ is the change in the rate of increase of climate system heat content in response to the forcing ΔF.

As energy budget estimates of ECS are directly grounded in basic physics and involve limited additional assumptions, unlike those from all other methods (including AOGCMs), they are particularly robust. The method does, however, rely on the use of reliable and reasonably well-constrained estimates of:

- changes in global mean total forcing
- TOA radiative imbalance (or its counterpart, climate system – very largely ocean – heat uptake)
- global mean temperature.

But providing that this is done, there seems little doubt that this approach should provide the most robust ECS estimates. Energy budget estimates in effect represent a gold standard.

The most important anthropogenic changes in the troposphere are:

---

[28] Given knowledge of what the forcing resulting from a doubling of atmospheric CO_2 concentrations ($F_{2\times \mathrm{CO}_2}$) is, ECS can be derived as follows: ECS = $F_{2\times \mathrm{CO}_2} \times \Delta T/(\Delta F - \Delta Q)$, where ΔT is the change in global temperature, ΔF the change in forcing and ΔQ the change in ocean heat uptake rate. While heat uptake by the atmosphere, ice and other non-ocean components of the Earth's climate system should in principle also be allowed for, they are almost negligible in relation to ocean heat uptake. Where the increase in forcing ΔF ends with, and mainly occurs during, an approximation to a ramp lasting 60–80 years, TCR may likewise be estimated as TCR = $F_{2\times \mathrm{CO}_2} \times \Delta T/\Delta F$. Note that estimation of forcings is assisted by some use of climate models. ECS and TCR estimates from energy budget methods can be affected by factors such as internal climate system variability, and assume that the physical relationships represented by these equations are stable over time and hold for a somewhat warmer climate than today's. But other observationally-based methods of estimating ECS and TCR are also affected by these and/or worse issues and generally involve substantially more assumptions.

[29] Section 10.8.1 of WGI. For clarity, the implicit multiplication sign in the second formula quoted has been made explicit.

- the increase in greenhouse gases, which leads to a positive forcing, implying warming
- the increase in aerosols, which by increasing reflection of incoming sunlight is thought, on balance, to lead to a negative forcing and therefore cooling.

Unlike the main greenhouse gases, the lifetime of aerosols in the troposphere is very short – only days to weeks – because they are removed from the atmosphere by rain.

The radiative effect of greenhouse gases is fairly well understood. However, the effect of aerosols is still rather uncertain. Deriving aerosol forcing from observations is difficult, and was not practicable prior to the development of suitable satellite instrumentation.

Gregory et al. compared the state of the climate between the two periods 1861–1900 and 1957–1994. Back in 2002 the authors had to use an estimate for aerosol forcing derived from climate models and not from observations,[30] so the study was not fully observationally-based. Their best estimate[31] for ECS came out at 6.1°C with a range of 1.6°C to infinity and was presented, truncated at 10°C, in a prominent figure in AR4.[32]

One of this report's authors (Lewis) worked his way through the Gregory et al. method. He discovered that Gregory's data for heat uptake in the oceans over 1957–1994 came from an erroneous dataset [33] that was corrected downwards in 2005, and that the total forcing change estimate Gregory had used was only half that used by NASA in their well-known GISS climate model. The combination of a low forcing change and a high ocean heat uptake change led to a high ECS estimate, with a very long upper tail. Use of the corrected ocean heat content (OHC) dataset and GISS model forcings reduced the ECS best estimate from 6.1°C to 1.8°C and gave a distribution that was much better constrained.[34]

---

[30] Although Gregory used an attribution method to scale the model-estimated aerosol forcing by reference to observations, the scaled model aerosol forcing change was over double that per the AR5 best estimate. Gregory also had to use a model estimate for ocean heat uptake during 1861–1900, but that factor was much smaller than aerosol forcing.

[31] All the best estimates given for ECS and TCR are medians.

[32] http://www.ipcc.ch/publications_and_data/ar4/wg1/en/figure-9-20.html. Reproduced in Figure 1.

[33] Levitus et al. (2000).

[34] http://www.judithcurry.com/2011/07/07/climate-sensitivity-follow-up/. Note that use of the AR5 forcing best estimates rather than GISS model forcings would have reduced the ECS best estimate even further.

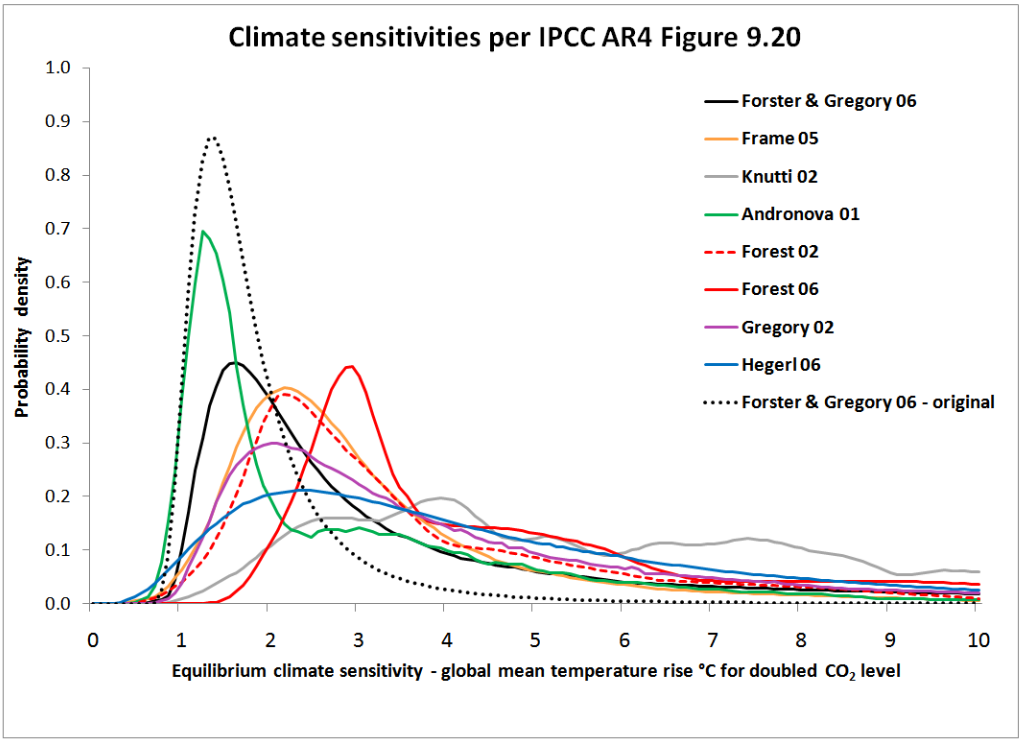

Figure 1: Replication of PDFs from Figure 9.20 in AR4 WGI

The dotted black line is an addition to those appearing in Figure 9.20 and shows the original results of Forster and Gregory (2006). All PDFs were scaled in AR4 to allocate all probability (one in total) between 0°C and 10°C.

How AR4 got sensitivity estimation wrong

Before discussing how climate sensitivity was treated in the AR5 report, it is useful to discuss how it was presented in AR4, published in 2007. Figure 9.20 in AR4 WGI gave estimated probability density functions (PDFs) for eight observationally based ECS studies (one a last-millennium palaeoclimate study[35]). These PDFs are replicated in Figure 1. The dotted black line is our addition to those appearing in Figure 9.20. We will comment briefly on the various studies to which these PDFs relate.

In Figure 1, most of the instrumental ECS estimates based on warming during a substantial part or all of the instrumental period involve comparing observations of climatic variables – usually more than just global mean surface

---

[35] Estimates from earlier periods, based on proxy data, are called 'palaeoclimate' estimates by the IPCC. As we discuss later, these involve greater uncertainties and do not narrow the ECS range derived from instrumental estimates.

temperature – at different times with multiple simulations by a simple or intermediate complexity climate model (model–observation comparison studies). Unlike AOGCMs, these models have adjustable parameters calibrated in terms of the model sensitivity and, often, other key climate system characteristics such as aerosol forcing and ocean heat uptake efficiency that are estimated alongside ECS. The closeness of the overall match between modelled and observed values at each setting of the model parameters[36] indicates how likely it is that setting represents the true values of those parameters.

Most of the studies adopted a Bayesian statistical approach. Bayesian analysis allows for the possibility that the researcher had some prior information as well, and the new data should merely update the prior estimate. Thus, rather than simply calculating a result, a weighting scheme is used to allow the new data to modify, rather than entirely replace, the prior understanding. The Bayesian approach is well suited to dealing with uncertain parameters and its use is not particularly controversial. But readers of scientific reports based on Bayesian analysis can easily overlook the fact that the choice of prior can significantly affect the result. If the chosen prior is itself controversial, this sharply downgrades the robustness of any findings.

Although the design of scientific studies may be informed by existing knowledge, once designed it is normal for their results only to reflect knowledge gained from the data used. Therefore, it is necessary for the prior estimate properly to represent, in mathematical terms, ignorance about climate sensitivity, not what the researcher believes about it before seeing the new data. In other words, the prior should be chosen to have minimal influence on the resulting ECS estimate: it should let the data speak for themselves. Hardly any of the priors used in AR4 and AR5 satisfied this requirement.

The Gregory et al. (2002) estimate has already been discussed. All the other studies, apart from Forster and Gregory (2006), were of the model–observation comparison type, one being a palaeoclimate study. That study, Hegerl et al. (2006), did not provide any useful constraint on ECS. Likewise, Knutti et al. (2002) found that their observational constraints did not enable a well constrained ECS estimate to be produced; they did place a lower limit of 1.2°C on ECS, but that figure was biased upwards by use of the same erroneous OHC

---

[36] How probable, given the uncertainties, obtaining the observed values was if the modelled values were correct.

dataset as Gregory et al. (2002). Moreover, both these studies used inappropriate 'uniform priors'[37] for ECS, biasing their estimates upwards. Since AR4, the use of a uniform prior for estimating ECS has been strongly criticised.[38] Three of the remaining studies used uniform priors both for ECS and ocean heat uptake efficiency, which biased their ECS estimates even more strongly upwards. Of those three studies, one[39] was also affected by statistical errors and poor experimental design, another[40] used erroneous OHC data, and the third[41] was affected by both these errors. Another study, Andronova and Schlesinger (2001) had unrealistic forcing assumptions and also appears to have been affected by an error in its computer code that substantially biased upwards its estimate for ECS.[42]

That leaves only one ECS study featured in AR4 to consider: Forster and Gregory (2006). Unlike all the other instrumental studies, it derived an estimate that was almost fully based on observations and did not have evident flaws such as faulty data or methodology. Forster and Gregory used satellite measurements of changes in the Earth's TOA energy imbalance and related those to changes in the global temperature. This gives a direct estimate of climate sensitivity, with little dependence on changes in aerosols.[43] In their original paper the results were fairly tightly constrained, meaning that the range of possible values for climate sensitivity was limited. Their best estimate was 1.6°C, their likely range was 1.2–2.5°C, and only 5% of their probability distribution lay above 4.1°C.

However, in AR4 the IPCC misrepresented the Forster and Gregory results by restating them on the inappropriate 'uniform prior in ECS' basis. The IPCC

---

[37] Equally weighting all possible values of a parameter within a wide (or unlimited) range initially, before weighting (at each such value) by the probability of obtaining the actual data. Typically, using uniform priors would be correct if the data used to estimate the parameters were linearly related to them. But where, as for these climate system parameters, the relationships are highly nonlinear then use of uniform priors greatly distorts estimation (unless the PDF is very narrow).

[38] Annan and Hargreaves (2011) wrote 'the uniform priors which have been widely used represent beliefs that in our opinion are extreme and difficult to justify'. Jewson (2013) stated that 'Flat [uniform] priors can almost never be justified…' and said that none of the discussed AR4 climate sensitivity studies that used them would give reliable estimates.

[39] Forest et al. (2006).

[40] Frame et al. (2005).

[41] Forest et al. (2002).

[42] See Ring et al. (2012).

[43] Although, as discussed in footnote 61, a study of this nature cannot readily distinguish the effects of cloud changes that are a response to surface temperature changes from independent natural internal cloud variability, Forster and Gregory concluded that this, in conjunction with the regression method used, was more likely to bias their ECS estimate upwards than downwards. In addition, the inclusion of 1992, the year when global temperature was significantly depressed by the eruption of Mount Pinatubo, is likely to have diluted the confounding effects of internal cloud variability.

curve – the solid black line in Figure 1 – is substantially skewed towards higher climate sensitivities and has a much fatter tail than the original results curve – shown in Figure 1 by the dotted black line. The top of the 'likely' range doubles, from 2.5 to 5.0°C, and the best estimate increases from 1.6 to 2.4°C.

AR4 contained several errors, including the well-publicised overestimate of the speed at which Himalayan glaciers might melt. However, the IPCC's defenders point out that such errors were inadvertent and inconsequential: they did not undermine the scientific basis of AR4. However, the distortion described here, which came about through the IPCC's unwarranted alteration of a peer-reviewed result, took place in the core scientific WGI report. This error was highly consequential,[44] since it involved the only good-quality empirical estimate of ECS cited by the IPCC, and the alteration substantially increased the apparent risk of high warming from increases in CO_2 concentration.

The message here is clear. AR4, published in 2007, could already have concluded from the instrumental evidence that the sensitivity of our real climate might well be lower than the 2°C lower limit that AOGCMs continued to suggest. Instead the IPCC AR4 authors actually raised the lower bound of the likely range for climate sensitivity from 1.5 to 2°C, while retaining a best estimate of 3°C. Note that the best estimate from the Forster and Gregory (2006) paper (1.6°C) fell *outside* this likely range.

The good news in AR5

Aerosols are – by their cooling effect on the climate – thought to have counteracted some of the warming from greenhouse gases. The effect of aerosols is the biggest uncertainty in estimates of total anthropogenic forcing. Knowledge of aerosols is therefore of crucial importance for estimates of climate sensitivity.

And this is where the AR5 report has some excellent news: its estimates of the cooling effect of aerosols are substantially lower than those in AR4. This in turn implies that sensitivity to greenhouse gases – both ECS and TCR – must be lower. The reasoning goes as follows.

---

[44] The IPCC does not accept that it was an 'error', on the grounds that the alteration of the Forster and Gregory (2006) PDF (although not its effect) was disclosed and is permissible under a subjective Bayesian philosophy (under which probability has no objective meaning). However, from a scientific viewpoint the altered statistical basis is indefensible.

We have had around 0.8°C of warming in the past one and a half centuries. The AR5 report presents evidence that this warming is almost entirely the result of a change in forcing, predominantly anthropogenic, with internal variability playing only a minor role. Since AR4 the concentration of greenhouse gases and the forcing they cause has increased, but although there has been little change in the emissions of aerosol-causing pollutants, the best estimate of aerosol forcing has come down substantially, from -1.3 W/m^2 in AR4 to -0.9 W/m^2 in AR5. This is purely a matter of what the IPCC called 'improved understanding'. As a result, the estimated net anthropogenic forcing has increased considerably, to 2.29 W/m^2. This is also reported in the SPM, which states (our emphasis):

> The total anthropogenic RF for 2011 relative to 1750 is 2.29 [1.13 to 3.33] Wm^{-2} (see Figure SPM.5), and it has increased more rapidly since 1970 than during prior decades. *The total anthropogenic RF best estimate for 2011 is 43% higher than that reported in AR4 for the year 2005.* This is caused by a combination of continued growth in most greenhouse gas concentrations and improved estimates of RF by aerosols indicating a weaker net cooling effect (negative RF). {8.5}

So, since AR4, estimates of the total anthropogenic radiative forcing are 43% higher and yet global temperatures remain almost unchanged.[45] Therefore, the same 0.8°C warming now has to be spread over considerably more units of forcing. Logically there is 30%[46] less warming per unit of forcing. But the warming per unit of forcing is a measure of climate sensitivity, in this case a measure close to TCR, and not ECS, since most of the increase in forcing has occurred over the last 60–70 years.

As AR5 states,[47] where there is an increase in forcing spread over a 70-year timescale, TCR is given by the following formula:

$$\text{TCR} = \frac{\text{Increase in temperature}}{\text{Increase in forcing}} \times \text{Forcing due to doubling CO}_2$$

The increase in temperature is almost the same now as for AR4. But the change in forcing is on the bottom of the fraction term: as it gets larger, the fraction must get smaller. Since the forcing due to a doubling of CO$_2$ factor ($F_{2\times\text{CO}_2}$) has

---

[45] The global surface temperature was marginally lower in 2012 than in 2007, as was the trailing pentadal mean temperature. The trailing decadal mean temperature was marginally higher.

[46] 100% $-$ 100%/143% = 30%. The change in estimated total forcing between AR4 and AR5 is predominantly in anthropogenic forcing.

[47] Section 10.8.1.

remained the same, the value of TCR must also fall as the increase in forcing rises. One can easily calculate what value of TCR is implied by the new AR5 forcing best estimates. AR5 estimates the forcing due to a doubling of CO_2 as 3.71 W/m^2. Therefore,

$$\mathrm{TCR} = \frac{0.8^\circ\mathrm{C}}{2.29\,\mathrm{Wm}^{-2}} \times 3.71\,\mathrm{Wm}^{-2} = 1.30^\circ\mathrm{C}$$

The implied value for TCR is just 1.3°C – much lower than the average value of between 1.8°C and 1.9°C from GCM estimates.

If we dig a little deeper into the full AR5 report the news gets even better. The best estimate the IPCC gives for total aerosol forcing is not fully based on observations. It is a composite of estimates derived from simulations by global climate models and from satellite observations. Six studies based on satellite-observations[48] with a mean best estimate of -0.78 W/m^2 were taken into account in deciding on the -0.9 W/m^2 AR5 composite best estimate of total aerosol forcing. So a best estimate for aerosol forcing based purely on satellite observations would be even smaller than the -0.9 W/m^2 that AR5 reports.

An ECS estimate based on AR5 estimates of changes in forcing and climate system heat uptake

We can use the robust energy-budget method of the Gregory 2002 paper to put all the AR5 data together, up to the most recent year.[49] We compare the periods 1859–1882 and 1995–2011. These two periods are the longest ones in, respectively, the earliest and latest parts of the instrumental period that were largely unaffected by major volcanic eruptions, the effects of which could affect ECS estimates. Another advantage of using the 1995–2011 period is that over this period the various datasets agree fairly well about changes in OHC.[50] We scale the aerosol component of AR5 total forcing estimates to match the recent satellite observation based mean of -0.78 W/m^2. Putting this all together gives a best estimate for ECS of 1.7°C, and of 1.30°C for TCR. Using AR5

---

[48] Out of nine satellite studies with best estimates ranging from -0.09 W/m^2 to -0.95 W/m^2 shown in Figure 7.19 of AR5.

[49] The HadCRUT4 v2 surface temperature record is used, as the other principal datasets do not extend back before 1880.

[50] To be conservative, we deduct only half (0.08 W/m^2) the allowance Gregory et al. (2002) made for ocean heat uptake in the early period.

forcing estimates without scaling aerosol forcing would give a best estimate for ECS of 1.76°C, and of 1.36°C for TCR.

So, based on the most up-to-date numbers from the IPCC report itself and using the most robust methodology, one arrives at observationally-based estimates for ECS and TCR that are very low: the ECS best estimate of 1.7°C is very close to the AR5 lower bound of 1.5°C and the TCR best estimate of 1.3°C is close to the AR5 lower bound of 1°C.

Evidence for low climate sensitivity piling up

Over the last two years several estimates of ECS have been published in the peer-reviewed literature using data from the instrumental period and methodology that appears satisfactory.[51] In particular, they incorporate observationally based aerosol forcing estimates. One of us (Lewis) was sole author of one of those studies, which is cited in several places in AR5 WGI. He is also a co-author of Otto et al. (2013), which is a notable paper because almost all of its other fifteen co-authors are also lead or coordinating lead authors of chapters of the AR5 WGI report that are relevant to the estimation of climate sensitivity.

In his own study Lewis comes up with a best estimate for climate sensitivity of 1.6°C, with a 'likely' range of 1.3–2.2°C.[52] The Otto et al. study gives a slightly higher sensitivity of 2.0°C , with a likely range of 1.5–2.8°C.[53]

All of the studies referred to were published in time to be included in AR5. All of them find best estimates for climate sensitivity of between 1.6°C and 2°C. These are shown in Table 2, which also gives the best estimates and ranges for ECS in AR4 and AR5, both the overall assessments and the estimates based on AOGCMs alone.

Ring et al. (2012), Aldrin et al. (2012), Lewis (2013) and Otto et al. (2013) are all based on observations encompassing the greater part of the instrumental period, and all use OHC as well as surface temperature measurements. Their

---

[51] Aldrin et al. (2012), Ring et al. (2012), Lewis (2013)and Otto et al. (2013).

[52] With non-aerosol forcing and observational surface temperature uncertainties incorporated

[53] Using data from the most recent decade considered, 2000–09, which arguably should provide the most reliable results. The Otto et al. study could have obtained a lower climate sensitivity best estimate had it used a different source of recent heat uptake data. See http://bishophill.squarespace.com/blog/2013/5/19/new-energy-budget-derived-estimates-of-climate-sensitivity-a.html. Using the heat uptake estimate from Loeb et al. (2012) would have resulted in a best estimate for ECS of 1.7°C, reducing to 1.6°C if the 2000–09 period were extended to 2012.

Table 2: Comparison of estimates for ECS from recent empirical studies that incorporate observationally-based aerosol forcing estimates, from models and from the IPCC reports

Study	**Best estimate**	**Likely range**	
		From	To
	°C	°C	°C
Ring et al. 2012 (using 4 surface temperature datasets)	1.80	1.4	2.0[a]
Aldrin et al. 2012 (main results)	1.76[b]	1.3	2.5
Lewis 2013 (preferred main results[c])	1.64	1.3	2.2
Otto et al. 2013 (2000s data)	2.00	1.5	2.8
Otto et al. 2013 (1970–2009 data)	1.91	1.3	3.0
Average of the above[d]	1.79[e]	1.3	2.4[f]
CMIP3 models (per AR4 Table 8.2)	3.20	2.1	4.4[g]
CMIP5 models (per AR5 Table 9.5)	2.89[h]	1.9	4.5[i]
IPCC AR4	3.00	2.0	4.5
IPCC AR5	None given	1.5	4.5

[a] Non-probabilistic range of 4 best estimates. [b] 1.53°C when using a uniform in 1/ECS prior distribution, which appears to be more objective than the main results uniform in ECS prior. [c] See footnote 52. [d] Giving a 50% weight to each of the two Otto 2013 estimates. [e] 1.73°C using the alternative Aldrin estimate based on the more objective prior for ECS. [f] 1.3–2.5°C based only on the probabilistic ranges, so excluding that for Ring et al. (2012). [g] 5–95% statistical fit range, but effectively downgraded in AR4 by incorporation within the 17–83% 'likely' range. It is not in fact clear that ECS ranges derived from an ensemble of climate models have a valid probabilistic interpretation. [h] The AR5 CMIP5 models have a mean ECS of 3.22°C but a median (best estimate) ECS of 2.89°C. For the AR4 CMIP3 models the median ECS was in line with the mean ECS. [i] 5–95% statistical fit range, downgraded as in AR4.

ranges make allowance for natural internal variability and other sources of uncertainty.[54]

Otto applied the energy budget equation to estimate ECS and TCR, using forcing, ocean heat uptake and global surface temperature data from four individual decades and for the full 40-year period 1970–2009, taking changes with respect to the period 1860–1879. Forcing was strongest during the final decade 2000–2009, which was also little affected by volcanic activity, resulting in arguably the most reliable estimates of ECS and TCR. The estimates based on the full 40-year period are probably next most reliable, since the smaller impact of internal variability over the longer period partially compensates for the weaker average forcing. In fact, best ECS estimates based on data for just the

---

[54] Apart from the simple non-probabilistic range for Ring et al. (2012).

1980s and just the 1990s are very similar to those based on data for 1970–2009, which demonstrates the robustness of the energy budget method.

The other three studies were of the model–observation comparison type. The climate models employed were run using many different combinations of ECS, aerosol forcing and ocean heat uptake efficiency values. Each of these three unknown parameters was then estimated by determining how well the model simulations matched historical observations of surface and ocean-layer temperatures. The models used temperatures that were resolved latitudinally, at least by hemisphere, and so were validly able to form their own estimates of aerosol forcing.[55] Otto et al. (2013), on the other hand, adjusted its total forcing data to reflect satellite observation based estimates of aerosol forcing. Thus all four studies used observationally-based aerosol forcing estimates.

All of these observational studies except Aldrin et al. (2012) used objective statistical methods, which should have led to their results accurately reflecting the data used, unlike many of the Bayesian observational studies featured in AR4 and AR5. Although Aldrin et al. (2012) used a uniform prior for ECS, it also gave alternative results using what appears to be a more objective Bayesian prior for ECS.[56] That best ECS estimate was 1.5°C, with a likely range of 1.2–2.0°C.

Heat going into the oceans

Around the initial publication of the AR5 report in late September 2013, media attention focused on what has been dubbed the 'hiatus' in global warming: the fact that for 15 years the global temperature has hardly risen. Several explanations have been suggested by the climate science community. A favourite is that heat accumulation has continued in the ocean – indeed accelerated since about 2000 – and that it is not therefore possible to say that the warming of the climate has stopped.

---

[55] Aerosols do not spread that far from where the pollution causing them is emitted, which is mostly in the northern hemisphere. Provided temperatures are resolved at least by hemisphere, these studies can reach well-constrained 'inverse' estimates of total aerosol forcing – typically similar to satellite-observation-based estimates of aerosol forcing. If, however, only global temperature data is used it is impossible to disentangle estimation of aerosol forcing from estimation of climate sensitivity and ocean heat uptake efficiency.

[56] A uniform in 1/ECS prior, the ECS estimate using which is similar to that using a non-Bayesian statistical method.

However, studies that use recent OHC data do take post-2000 heat inflow into the oceans into account. For example, among the ECS estimates in Otto et al. (2013), the one based on the most recent decade's data is actually higher than the ones based on the 1990s or the 1970–2009 period. While this may seem counterintuitive, the reason is that – according to the 0–700m ocean layer heat content dataset[57] that it used (which was also used in AR5) – ocean heat uptake was much higher in the first decade of this century than in the previous decade.

Other OHC datasets and estimates from satellite radiative imbalance based studies do not show an acceleration of heat uptake to such high levels in the 2000s. If Otto et al. (2013) had used one of these instead, its ECS estimate based on 2000s data would have been in the range 1.7–1.9°C, depending on which particular dataset or estimate was used.

Thus these ECS estimates take the more-heat-going-into-the-ocean explanation for the near standstill in global surface temperature fully into account: although more heat going into the deep oceans might be an explanation for the slowdown of the warming at the surface, it does not materially change our estimates of ECS. These are still far lower than the best estimate of 3°C that has been prevalent over the last thirty years. The hiatus does, however, decrease estimates for the TCR, which is thought to be a more policy-relevant measure.

Best estimate

In order to calculate a 'best observational' estimate of ECS, it is reasonable to take a simple average of the different observationally-based estimates in Table 2, since all these studies use similar observational data. This gives a best estimate for ECS of 1.75°C and a likely range of about 1.3–2.4°C.[58] A 1.75°C best estimate is supported by the energy budget estimate based on AR5 data, given above, even without scaling the AR5 aerosol forcing estimate to match the satellite observations.

However, recognising that uncertainty arising from internal variability, measurement and model error may be greater than that allowed for, and predominantly affects the upper bound of the range, we conservatively assess the

---

[57] An update of that in Domingues et al. (2008).

[58] Based on the best estimate for Aldrin et al. (2012) employing the more objective uniform in 1/ECS prior, marginally rounding up the resulting average best estimate, and using only the probabilistic ranges.

likely range to be 1.25–3.0°C, which encompasses the full 17–83% probabilistic uncertainty range for each of the observationally-based estimates cited.

Now compare those figures with both the best estimate and the range in IPCC AR4 and AR5 (see also Table 1 for the longer historical evolution of the range). The 'best observational' ECS estimate of 1.75°C is more than 40% lower than both the best estimate in AR4 of 3°C and the 3.2°C average ECS of AOGCMs used in AR5. At least as importantly, the top of the likely range for ECS of 3.0°C is a third lower than that given in AR5 (4.5°C), even after making it much more conservative than is implied by averaging the ranges for each of the observational estimates.

Poor estimates obscure the issue

Of course, the four studies included in Table 2 represent only a part of one line of evidence cited in AR5 as to the value of ECS. Box 12.2, Figure 1 in AR5, reproduced below as Figure 2, shows the ranges from many sources, categorised by the line of evidence. From just looking at this figure it seems possible to understand why AR5 did not give a best estimate: the best estimates for ECS are not clustered around a single value and many of them are outside the uncertainty ranges – here 5–95% ranges, not 17–83% ('likely') – from other studies. The AR5 authors might not have wanted to decide that some studies are better than others or to adjudicate between observational and model based lines of evidence, but in our opinion using expert knowledge to weigh different sources of evidence is exactly what an assessment is all about. Here we present reasoned arguments for a different assessment to that in AR5.

We will therefore discuss the estimates in Figure 2 in some detail, showing why little weight should be put on those estimates that are inconsistent with the likely ranges for the 'best observational' studies in Table 2, either because of some identified serious shortcoming in their derivation or because they use a method upon which AR5 itself casts doubt. In this connection, we will accept the conclusion in AR5[59] that estimates of ECS based on:

- past climate states very different from today
- timescales different than those relevant for climate stabilization (e.g. climate response to volcanic eruptions)

---

[59] Section 12.5.3.

- forcings other than greenhouses gases (e.g. volcanic eruptions or solar forcing)

may differ from the climate sensitivity based on the climate feedbacks of the Earth system today. Accordingly, so far as observational estimates of ECS are concerned we concur with the AR5 authors that reliance should primarily be placed on instrumental estimates based on warming during a substantial part or all of the period since 1850.

Instrumental estimates

The unlabelled ranges in Figure 2 are for studies cited in AR4, all of which have already been discussed and reasons given as to why they are, with one exception, unsatisfactory. The exception is the original Forster and Gregory (2006) range (the unlabelled solid mauve bar fourth up from the bottom of the Instrumental section in Figure 2), the 17–83% 'likely' segment of which is in line with the 'Average' range in Table 2.

The labelled, AR5, ranges for instrumental estimates not included in Table 2 can be dealt with as follows:[60] Bender et al. (2010) is based on the response to volcanic forcing. The Lindzen and Choi (2011) and Murphy et al. (2009) studies are based on short-term variations in TOA radiative fluxes as measured by satellite.[61] Both of these methods are regarded in AR5 as not being satisfactory ways of constraining ECS.[62]

---

[60] Only the solid line ranges in Figure 2 for Aldrin et al. (2012) and Lewis (2013), which represent the (primary/preferred) main results, are reflected in Table 2 (for Lewis (2013), after incorporating non-aerosol forcing and observational surface temperature uncertainties); the other ranges are on different bases. The higher, dashed, range for Lewis (2013) is based on a shorter data period and less satisfactory diagnostics, matching those in Forest et al. (2006), than the solid line preferred main results range, and should be disregarded since its best-fit parameter settings are known to produce excessive simulated warming. For Aldrin et al. (2012), the lower, dashed, range is the one using the uniform-in-1/ECS prior discussed earlier, whilst the dash-dotted range should be disregarded as it is biased by use of an excessively negative aerosol forcing prior.

[61] A major uncertainty in such studies arises from the difficulty of distinguishing the effects of cloud changes that are a response to surface temperature changes (and thus constitute feedbacks) from independent natural internal cloud variability (which acts as a forcing). Lindzen and Choi (2011) seek to distinguish these factors by using lagged regression; their results indicate ECS is very low. This is a clever approach, but it is not yet clear how robust the results are. Murphy et al. (2009), using unlagged regression, reach an estimate for ECS that is high and poorly constrained. Forster and Gregory (2006), obtained a low (1.6°C), well-constrained ECS estimate when using unlagged regression and almost the same data from which Murphy et al. obtained a much higher estimate. (See footnote 43 regarding a possible explanation for this difference.)

[62] Sections 10.8.2.2 and 10.8.2.3 of AR5 give detailed reasons for doubting the usefulness of ECS estimates based on these methods.

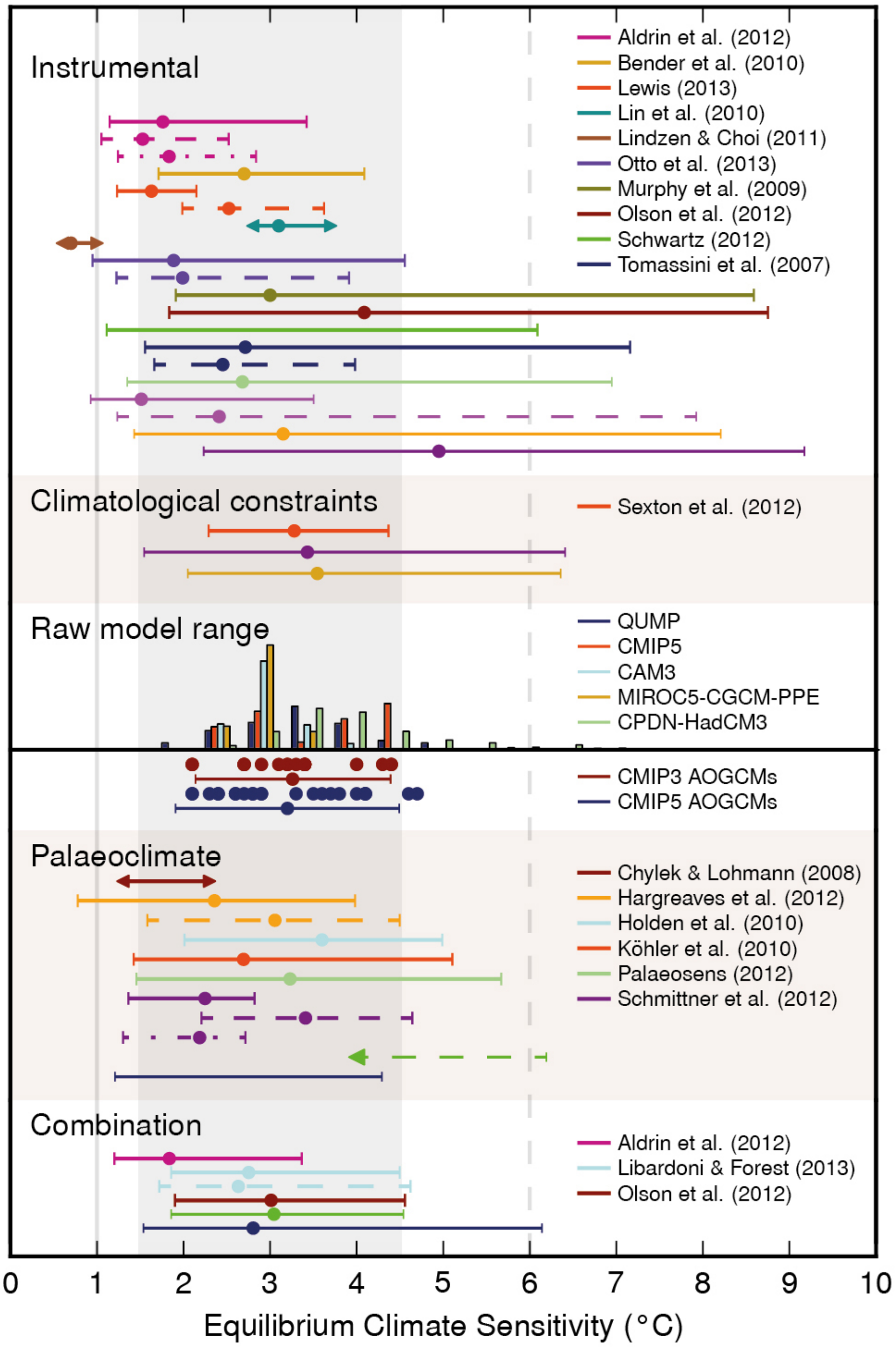

Figure 2: Reproduction of Box 12.2, Figure 1 from AR5

Bars show 5–95% uncertainty ranges for ECS, with the best estimates marked by dots. Actual ECS values are given for CMIP3 and CMIP5 AOGCMs. Unlabelled (thin bars) ranges relate to studies cited in AR4.

The remaining AR5 studies, Lin et al. (2010), Olson et al. (2012), Schwartz (2012) and Tomassini et al. (2007) all have identified shortcomings, set out in the Appendix, that make their estimates of ECS unsatisfactory.[63]

Climatological constraints

Climatological constraint studies reflect how good one or more GCMs are at simulating various aspects of the recent climate as key model parameters are perturbed so as to produce different model behaviour and hence ECS values. The assumption is that from such an exercise it is possible to infer what range of ECS is most likely. The method presupposes that by perturbing its parameters a GCM will be able to explore all combinations of reasonably possible aerosol forcing and ECS values. Even if that is so, it is unclear whether comparisons with observations of aspects of recent climate, as opposed to climate change, can produce reliable ECS ranges.

The Sexton et al. (2012) study involved perturbing parameters of the UK Met Office HadCM3 climate model, generating different ECS values.[64] However, due to structural rigidities in the HadCM3 model, no matter what parameter combination is used, when low ECS values are achieved by the model, its aerosol forcing becomes very highly negative[65] – a combination ruled out by the observational data. The Sexton study was unable to investigate the combination of low-to-moderate ECS and low-to-moderately negative aerosol forcing – the region favoured by the observational data. It is thus unsurprising that the study rules out low ECS values. Its ECS estimate very largely reflects the characteristics of the HadCM3 model rather than the observations.

The two unlabelled AR4 studies, although differing in detail from Sexton et al. (2012), also use the HadCM3 model. Therefore, they will likewise have been unable to investigate the combination of low-to-moderate ECS and low-to-moderately negative aerosol forcing. Moreover, it appears that both these studies barely sampled ECS values below 2°C. These two shortcomings can be

---

[63] In the case of Schwartz (2012) the criticism relates only to the section of its range for ECS that exceeds 3°C.

[64] Emulation is used to extrapolate to ECS values of 2°C and below, since HadCM3 has structural rigidities that make it unable to exhibit low ECS values no matter how its parameters are adjusted.

[65] See Box 1 in the document at http://niclewis.files.wordpress.com/2013/09/metoffice_response2g.pdf. The Sexton et al. (2012) study is identical to the first stages of the Harris et al. (2013) study that it discusses, and the Harris et al. near-final posterior region in Box 1 Figure B.1 corresponds to the final results of the Sexton et al. (2012) study.

expected to have strongly biased upwards estimation of ECS in the two AR4 studies.

In summary, since they strongly reflect characteristics of particular GCMs, and only weakly reflect observational evidence, the climatological constraint studies are of little or no value as estimates of ECS.

Raw model range

As the conflict between observational and AOGCM estimates of ECS is a central issue, and AR5 also cites a related line of evidence referred to as feedback analysis (the analysis of climate feedbacks simulated by AOGCMs), it is appropriate to discuss in some detail feedbacks and ECS in climate models.

It is almost universally accepted that by itself the equilibrium warming effect of a doubling of the CO_2 concentration – the amount global mean temperature needs to rise for the resulting increase in emitted black-body (Planck) radiation to offset the increase in CO_2 forcing – is slightly more than 1°C. Why then do models have an average ECS of 3°C? This is due to so-called 'positive feedbacks' caused by the initial increase in surface temperature. Positive feedbacks amplify the warming effect of CO_2. The main climate feedbacks in the models are water vapour, lapse rate, cloud and albedo feedbacks. Water vapour feedback is very strongly positive: a warmer atmosphere can hold more water vapour, which itself is a powerful greenhouse gas. Water vapour feedback is partially cancelled by the related negative lapse rate feedback: with moister air, temperature decreases more slowly with altitude. Less snow and sea ice cover as temperatures rise makes the Earth reflect less sunshine, hence the albedo feedback is positive, albeit fairly weak. Together, these three feedbacks, when aggregated with the basic Planck radiation change, imply an ECS of around 2°C.[66] The excess of model ECS over 2°C comes primarily from positive cloud feedbacks and adjustments, with non-linearities and/or climate state dependency also having a significant impact in some cases.

But clouds are a big headache for the modellers. It is very difficult to simulate them, let alone to predict how they will change in the future. Observational evidence for cloud feedback being positive rather than negative is weak, at

---

[66] Soden and Held (2006); Table 9.5 of AR5 WGI. ECS is estimated as $F_{2\times CO_2}/\alpha$ in accordance with Section 10.8.1 of AR5 WGI, but in this case with only the Planck, water vapour, lapse rate and albedo feedbacks included in α; $F_{2\times CO_2}$ is taken as 3.71 W/m^2.

best.[67] Even the observational evidence for the modelled strength of water vapour feedback is rather thin on climatic time scales.[68]

There is no knob for climate sensitivity as such in global climate models, but there are many adjustable parameters affecting the treatment of processes (such as those involving clouds) that GCMs do not calculate from basic physics. Climate sensitivities exhibited by models that produce realistic simulated climates, and changes in climatic variables over the instrumental period, are assumed to be representative of real-world climate sensitivity. However, there is no scientific basis for this assumption. An experienced team of climate modellers has written that many combinations of model parameters can produce good simulations of the current climate but substantially different climate sensitivities.[69] They also say that a good match between AOGCM simulations and observed twentieth century changes in global temperature – a very common test, cited approvingly in the AR4 IPCC report as proving model skill – actually proves little. Models with a climate sensitivity of 3°C can roughly match the historical record for the global temperature increase in the twentieth century, but only by using aerosol forcing values that are larger than observations indicate is the case, by underestimating positive forcings, by putting too much heat into the oceans and/or by having strong non-linearities or climate state dependency.[70]

If there were broad agreement between AOGCMs as to the sign and – within, say, a factor of two – the magnitude of all significant feedbacks, and as to their spatial dependencies, and nonlinearities and climate state dependencies were qualitatively similar across AOGCMs, then it would be reasonable to place significant weight on AOGCM-based evidence about climate sensitivity. However, despite model development being closely informed by diverse observations, that is not the case. So we think one should disregard AOGCM-based evidence about climate sensitivity – everything shown in the 'Raw model range' section. Being only tenuously grounded in observations, it is unclear to what extent raw model ECS values qualify as scientific evidence at all. Likewise, the related evidence as to ECS based on analyses of feedbacks in models (discussed in Chapter 12 of AR5, but not featured in Box 12.2, Figure 1) should be disregarded both because it is not evident that all significant feedback processes are included in models and because a critical part – evidence as to cloud feedbacks – is very unsatisfactory.

---

[67] Section 7.2.5.7 of AR5.

[68] VonderHaar et al. (2012).

[69] Forest et al. (2008).

[70] Based on the best estimates of forcing per AR5 and best estimates of OHU using a range of OHC datasets, this is implied by the results of Otto et al. (2013), which follow from conservation of energy.

Palaeoclimate

Proxy-based palaeoclimate studies estimate climate sensitivity by using the climate records of the more distant past (the last millennium, the period back to the last glacial maximum, or even back to millions of years ago). However, in 2007 the AR4 report concluded (Box 10.2) that uncertainties in last glacial maximum studies were too great for them to be regarded as providing primary evidence as to ECS, and the one ECS range it gave from a last-millennium proxy-based study only very weakly constrained ECS.[71] No results from other last-millennium studies were included in Box 12.2, Figure 1.

AR5 also discusses palaeoclimate estimates. It says of a recent review article:[72] 'They estimate a 95% range of 1.1°C–7.0°C, largely based on the past 800,000 years. However, uncertainties in palaeoclimate estimates of ECS are likely to be larger than from the instrumental record, for example, due to changes in feedbacks between different climatic states.' With such wide uncertainty ranges, palaeoclimate ECS estimates contain little information.

So AR5 takes the view that palaeoclimate ECS estimates based on past climate states that are very different from today may not be representative of the current state of the climate system, and are likely to provide less good constraints on ECS than do instrumental studies. That is broadly what AR4 said – it regarded palaeoclimate estimates as useful supporting information rather than primary evidence as to the level of ECS. Accordingly, little weight can be put on the palaeoclimate estimates.

Combination

Combination studies are estimates based on combining information from different methods. Of the labelled studies shown in AR5, the Aldrin et al. (2012) ECS estimate is little different from its main instrumental estimate, while the Libardoni and Forest (2013) and Olson et al. (2012) papers and the unlabelled AR4 studies have serious shortcomings and their combination estimates of ECS are all unsatisfactory (see Appendix).

---

[71] Hegerl et al. (2006). As discussed above, this study used an inappropriate uniform prior for ECS, biasing its ECS estimate upwards.

[72] Paleosens Members (2012), discussed in Section 10.8.2.4 of AR5.

Instrumental estimates are superior

So, to conclude, we think that of the three main approaches for estimating ECS available today (instrumental observation based, palaeoclimate proxy-observation based, and GCM simulation/feedback analysis based), instrumental estimates – in particular, those based on warming over a substantial period extending to the twenty-first century – are superior by far. Observationally-based estimates give the best indication of how our current climate has actually been reacting to the increase in greenhouse gases. Our view as to which form of observational study provides the most reliable method of estimating ECS is strongly supported by what Chapter 12 of AR5 has to say:[73]

> Equilibrium climate sensitivity undoubtedly remains a key quantity, useful to relate a change in greenhouse gases or other forcings to a global temperature change. But the above caveats imply that estimates based on past climate states very different from today, estimates based on timescales different than those relevant for climate stabilization (e.g., estimates based on climate response to volcanic eruptions), or based on forcings other than greenhouses gases (e.g., spatially non-uniform land cover changes, volcanic eruptions or solar forcing) may differ from the climate sensitivity measuring the climate feedbacks of the Earth system today, and this measure, in turn, may be slightly different from the sensitivity of the Earth in a much warmer state on timescales of millennia.

Furthermore, we have identified substantial shortcomings, rendering them unreliable, in every single one of the observational estimates for ECS cited in AR5 that are based on warming during the instrumental period other than those included in Table 2, the latter having best estimates in the range 1.6–2°C. Indeed, where a study uses forcing and heat uptake estimates that are consistent with those in AR5, that is almost bound to be the case on conservation of energy grounds.[74]

On our reading of AR5, the IPCC scientists largely agreed with our analysis of the observational evidence about ECS.[75] However, they were stuck with the

---

[73] Section 12.5.3.

[74] By comparison with the results of energy budget analyses. The only exception would be if a study produced its own properly constrained inverse estimate of the rather uncertain aerosol forcing, which exceeded the AR5 best estimate thereof. None of the instrumental studies did so.

[75] The AR5 Technical Summary justifies the reduction made in the lower bound of the likely range for ECS as a reflection of the evidence from new studies of observed temperature change, using the extended records in atmosphere and ocean, which suggest a best fit to the observed surface and ocean warming for ECS values in the lower part of the likely range.

ECS range corresponding to CMIP5 models and the lines of evidence as to ECS derived from them and from the observationally-based estimates that we criticise. In our view, the conflict between ECS estimates based on new studies of observed temperature change over the instrumental period and those based on models was probably the most important factor in the IPCC authors' decision not to give a best estimate for ECS this time.

The conflict between the best estimate of ECS implied by the latest observational evidence and that based on the CMIP5 models presented the IPCC authors with a dilemma. Large parts of the IPCC reports are built around the computer model simulations. Almost all the projections of future climate change are based on them, and a complete chapter is devoted to model performance.

Stating in the SPM that the best observationally-based estimates of climate sensitivity now indicate a value of only 1.5–2°C would come very close to an admission that most of the CMIP5 GCMs, at least, substantially overestimate ECS, which – since the projected warming towards the end of this century is strongly correlated with ECS in the GCMs[76] – would imply that policy makers should not place reliance on longer-term model-based climate projections.

It appears that the IPCC authors may have decided to resolve this dilemma by reducing the lower bound of ECS to 1.5°C and omitting a best estimate completely. By doing this they went some way to reflect the new, lower estimates that have been published recently in the literature. Now of course the IPCC scientists are quite entitled to reach a different conclusion from us as to whether much weight should be placed on model-based ECS estimates. However they failed to discuss this issue clearly in the SPM, thereby leaving policymakers in the dark.

The IPCC could have said 'there are two main methods to estimate ECS and one – based on observations – indicates, using the best quality data and sound methodology, that ECS is most probably 2°C or slightly less. The other – based on models – indicates it is about 3°C.' That would have been a step in the right direction because at least policymakers would have been alerted that model-based estimates are starting to deviate substantially from observational estimates.

In our review comments one of us (Lewis) advised the AR5 authors to show best estimates for both (instrumental and model based) methods separately:

---

[76] See also footnote 4.

> It is very important to keep the range for ECS estimated from observations – particularly instrumental observations, which as well as being more accurate also relate to the current climatic conditions – separate from that derived from AOGCM simulations. AOGCMs may, directly or indirectly, use forcing or other inputs that are not consistent with the best current observational evidence. That is a particular concern in relation to aerosol forcing, and also ocean effective vertical diffusivity, either or both of which may be substantially overestimated in AOGCMs, leading to excessive levels of ECS nevertheless producing realistic simulations of past warming. For instance, the NASA GISS global climate models now assume recent (2010) total aerosol forcing of -2.42 W/m^2 (http://data.giss.nasa.gov/modelforce/RadF.txt), over three times the best purely observational best estimate per AR5 of -0.73 W/m^2.

Note that this comment was written in November 2012 during the review of the Second Order Draft. Since then other studies (Ring et al. 2012; Lewis 2013; Otto et al. 2013) have been published and it is now possible to give the much better constrained likely range for ECS of 1.25–3.0°C based on (but more conservative than) that derived in Table 2. The IPCC could have additionally given a 'best observational' estimate of 1.75°C or (taking into account higher estimates from other instrumental studies) 2°C. If the IPCC had made that change – in line with the best quality scientific evidence available– it would have been picked up by all the major news outlets in the world as one of the major, if not the major, outcomes of the report. And rightly so.

Unsatisfactory treatment in AR5

AR5 states in a footnote in the SPM that no best estimate for ECS can be given this time, because of a lack of agreement on values across assessed lines of evidence and studies. Explaining such an important decision only in a footnote is unsatisfactory. Policymakers should have been given a full explanation.

Several of the underlying chapters deal with this issue: the Technical Summary and Chapters 9, 10 and 12. On page 84 of the Technical Summary the reduction in the lower bound of the ECS 'likely' range is discussed (our emphasis):

> This change reflects the evidence from new studies of observed temperature change, using the extended records in atmosphere and ocean. *These studies suggest a best fit to the observed surface and ocean warming for ECS values in the lower part of the likely range.* Note that these studies are not purely observational, because they require an estimate of the response to radiative forcing from models. In addition, the uncertainty in ocean heat uptake remains substantial.

Here AR5 quite openly admits that these new (observationally-based) studies have best estimates close to the lower bound of 1.5°C. The new studies' probabilistic ECS ranges allow for uncertainty in ocean heat uptake, which AR5 in fact estimates to be much less significant than uncertainty in aerosol forcing.

However, when the approved SPM was published even the full report was almost silent about the lack of a best estimate for ECS. The Chapter 10 section about climate sensitivity ends with the following statement:

> In conclusion, estimates of the Equilibrium Climate Sensitivity…based on multiple and partly independent lines of evidence from observed climate change, including estimates using longer records of surface temperature change and new palaeoclimatic evidence, indicate that there is high confidence that ECS is extremely unlikely less than 1°C and medium confidence that the ECS is likely between 1.5°C and 4.5°C and very unlikely greater than 6°C. They complement the evaluation in Chapter 9 and support the overall assessment in Chapter 12 that concludes between all lines of evidence with high confidence that ECS is likely in the range 1.5°C to 4.5°C. Earth system feedbacks can lead to different, probably larger, warming than indicated by ECS on very long timescales.

Again, a best estimate for ECS was not even mentioned. While the lack of any detailed information in the SPM about the decision not to give a best estimate for ECS might be argued as reflecting space limitations, the silence in the relevant chapters (10 and 12) of the full report, where one would have expected a detailed explanation about this decision, is more surprising.

However, in the final report published in January 2014 a paragraph was inserted into the Technical Summary discussing the fact that no best estimate for ECS can now be given.[77] This is quite surprising. Edits at this very late stage are meant to correct errors.[78] The section explaining no best estimate for ECS is not the correction of an error, it is just new text. This new paragraph, revealed long after governments approved the SPM, says this:

> In contrast to AR4, no best estimate for ECS is given because of a lack of agreement on the best estimate across lines of evidence and studies and an improved understanding of the uncertainties in estimates based

---

[77] The IPCC provided a long list of substantive edits that have been made after the final draft of the report: http://www.ipcc.ch/report/ar5/wg1/docs/review/WG1AR5_SubstantiveEditsList_All_Final.pdf.

[78] The front sheet to the accepted final draft of the AR5 WGI report published on 30 September 2013 stated: 'Before publication the Report will undergo final copyediting as well as any error correction as necessary, consistent with the IPCC Protocol for Addressing Possible Errors.'

> on the observed warming. Climate models with ECS values in the upper part of the likely range show very good agreement with observed climatology, whereas estimates derived from observed climate change tend to best fit the observed surface and ocean warming for ECS values in the lower part of the likely range. In estimates based on the observed warming the most likely value is sensitive to observational and model uncertainties, internal climate variability and to assumptions about the prior distribution of ECS. In addition, 'best estimate' and 'most likely value' are defined in various ways in different studies.

So here AR5 finally gives some additional explanations. The reader could, however, be wrongfooted by the remark that climate models (AOGCMs) with high ECS values are in good agreement with 'observed climatology'. This simply means they simulate certain properties of the current climate quite well; it does not mean they simulate global warming well. The authors then caveat the observational estimates by mentioning various issues that, where significant, are normally taken account of in sound studies.

Models overestimate recent warming

Much of the information in the AR5 report is based on simulations by the latest generation of AOGCMs (the so-called CMIP5 models). More than twenty groups around the world performed special runs of their climate models for the AR5 report. These models simulate the warming over the past 150 or so years and the simulations are then continued to give projections of future climate change, using different scenarios of future greenhouse gas concentrations. These projections are important for policy purposes. They give an idea of how much warming future emission paths will give rise to and therefore how ambitious mitigation policies have to be to achieve the targets set for maximum rises in global temperature.

The virtual climates in the GCMs turn out to be much more sensitive to CO_2 and other greenhouse gases than the best observational evidence indicates the real climate is (see Table 2). The CMIP5 models ultimately warm on average about 3.2°C[79] when the concentration of CO_2 is doubled. This is approaching twice the level suggested by the best observational studies. By not giving a best estimate the IPCC avoided having to reveal this difference between observational and model-based estimates of climate sensitivity.

---

[79] The CMIP5 mean ECS value rather is quoted here rather than the median, since AR5 shows means rather than medians in its CMIP5 multimodel based projections of future warming.

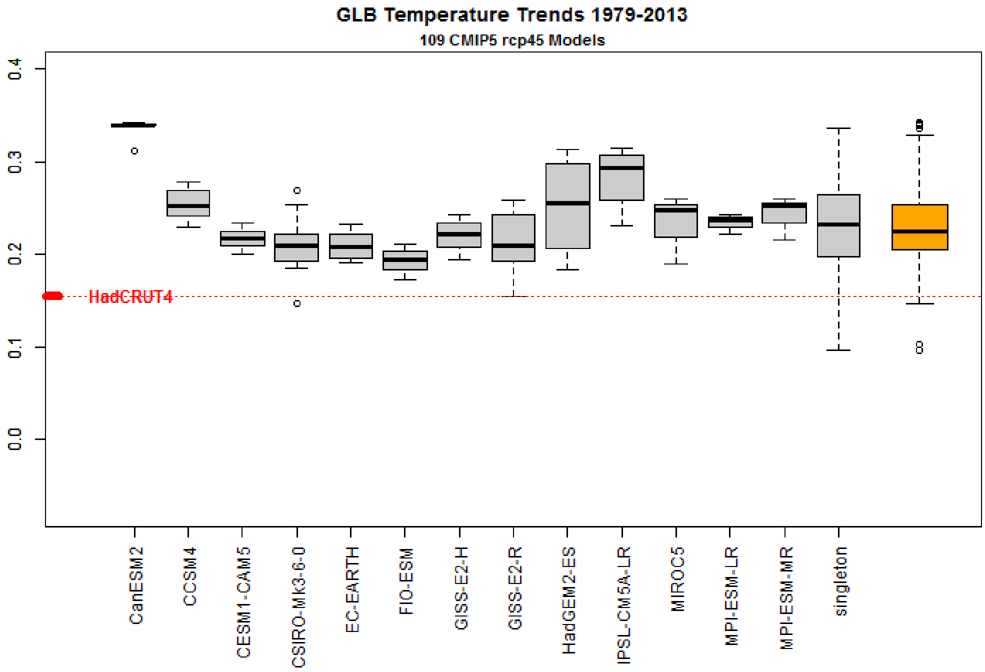

Figure 3: Modelled versus observed decadal global surface temperature trend 1979–2013

Temperature trends in °C/decade. Virtually all model climates warmed much faster than the real climate over the last 35 years. 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 default settings in the R boxplot function have been used; the end of the boxes represent the 25th and 75th percentiles. The red dotted line shows the actual trend in global surface temperature over the same period per the HadCRUT4 observational dataset.

A lot of the recent public attention has been focussed on the slowdown of global warming in the last fifteen years, which the climate models failed to predict. Defenders of the models tend to admit that models have difficulties with natural fluctuations in the climate that last for 10 to 15 years. However, the situation is much worse. Virtually all the models that the IPCC uses in its report have been running too hot over periods as long as 35 years, long enough to judge them on a climatic timescale (see Figure 3).[80]

---

[80] The 1979–2013 observed global temperature trends from the three datasets used in AR5 are very similar; the HadCRUT4 trend shown is the middle of the three. Several bloggers have recently shown excessive model warming over various periods, for example Steve McIntyre http://climateaudit.org/2013/09/24/two-minutes-to-midnight/, Lucia Liljegren http://rankexploits.com/musings/2013/

Note that Figure 3 is from a blog article. Nowhere in AR5 is a similar graph available. The one that comes closest is Figure 1 from Box 9.2, reproduced here as Figure 4.

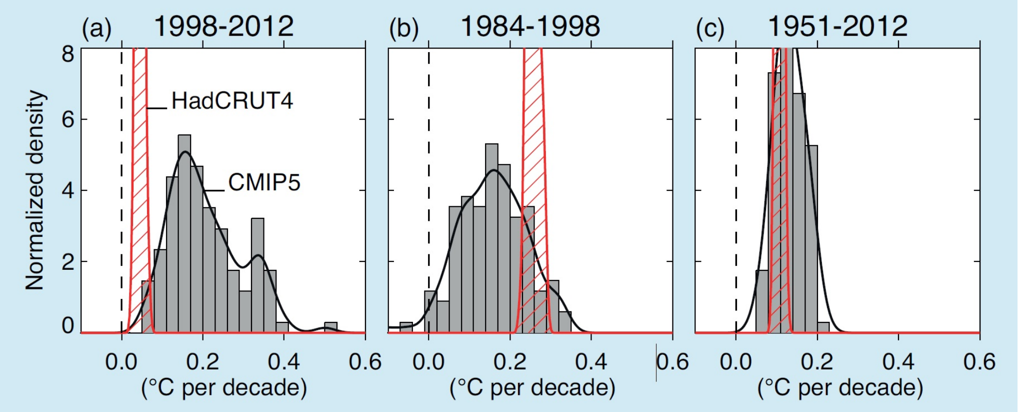

Figure 4: Frequency distribution of trends in global mean surface temperature from 114 CMIP5 model runs

Model runs are grey bars for the periods (a) 1998–2012, (b) 1984–1998, (c)1951–2012. The comparison is with the uncertainty range for the observed trend per the HadCRUT4 dataset (red, hatched) over the same periods. Reproduced from AR5, Box 9.2, Figure 1.

In this figure the IPCC attempts to show that the recent hiatus is more to do with choosing the hot El Niño year 1998 as a starting point. Panel (a) shows that CMIP5 models overestimate the HadCrut4 global temperature trend since 1998. However in panel (b) one can see that models tend to underestimate the observations in the period 1984–1998. So the message is: if you look at short periods of 15 years the models are sometimes too hot and sometimes too cold. Panel (c) then suggests models are performing well on a longer timescale, in this case 60 years. That is not surprising, since models are likely to have been tuned so that they provide a reasonable match to the global surface temperature rise over the historical simulation period, most of which occurred after 1950. The discrepancy between models and observations over the last 35 years is conveniently not shown. This period is long enough to be relevant for climate.

So models overestimate the warming of the real climate in the last 35 years by 50%. And the same models have ECS and TCR values that are considerably

---

leaked-spm-ar5-multi-decadal-trends/ and Roger Pielke Jr. http://rogerpielkejr.blogspot.nl/2013/09/global-temperature-trends-and-ipcc.html. A recent commentary in *Nature Climate Change* by Fyfe et al. (2013) reached similar conclusions.

higher than estimates based on observations indicate. Both these important observations were not made explicitly by the IPCC in AR5.

Transient climate response in AR5

So far we have mainly discussed the scientific evidence for estimates of ECS. But it takes centuries to millennia before the climate system reaches a new state of equilibrium and therefore climate scientists tend to regard TCR as more policy relevant. Andrews and Allen (2008) wrote that 'TCR is also the key determinant of climate change during the 21st century'.[81]

AR5 showed in its Figure 10.20(a), reproduced as Figure 5 here, a range of observationally-based TCR estimates. One of us (Lewis) has written a critical analysis of many of these TCR studies.[82] It finds serious fault with all the studies other than Gillett et al. (2013), Otto et al. (2013) and Schwartz (2012). The analysis notes that the individual CMIP5 model observationally-constrained TCRs shown in a figure in the Gillett et al. (2013) study imply a best estimate for TCR of 1.4°C, with a 5–95% range of 0.8–2.0°C.[83] The Otto et al. (2013) TCR range of 0.9–2.0°C using 2000–2009 data has a best estimate of 1.3°C, compared with slightly over 1.35°C using the lower signal-to-noise ratio 1970–2009 data. The Schwartz (2012) range is marginally lower at 0.85–1.9°C, with a best estimate of 1.3°C. A best estimate for TCR of 1.3°C was also derived earlier in this report (see page 25) from an energy budget analysis using a 1995–2011 final period and AR5 forcing estimates.[84] There is a detailed discussion of that estimate and of the observational TCR estimates cited in Figure 10.20(a) of AR5 in a post at the *Climate Audit* blog.[85]

The 'likely' range for TCR given in AR5 is 1–2.5°C, with TCR 'extremely unlikely' to exceed 3°C. That represents only a marginal reduction compared with AR4, where TCR was assessed to be 'very likely' to lie in the range 1–3°C. No best

---

[81] However, whilst this is almost true by definition for the real climate system, in CMIP5 models it is not clear that TCR provides a better guide than ECS to projected warming towards the end of this century, when both are scaled optimally.

[82] See http://niclewis.files.wordpress.com/2013/11/ar5_tcr_estimates2.pdf.

[83] As well as being based on regressions on a model-by-model basis rather than a single regression incorporating all models at once, this range does not allow for as wide a range of uncertainties as the range shown in Figure 5.

[84] With aerosol forcing scaled to match the mean of the satellite observation estimates used in forming the AR5 range for estimated aerosol forcing; without such scaling the TCR best estimate is 1.36°C.

[85] http://climateaudit.org/2013/12/09/does-the-observational-evidence-in-ar5-support-itsthe-cmip5-models-tcr-ranges/

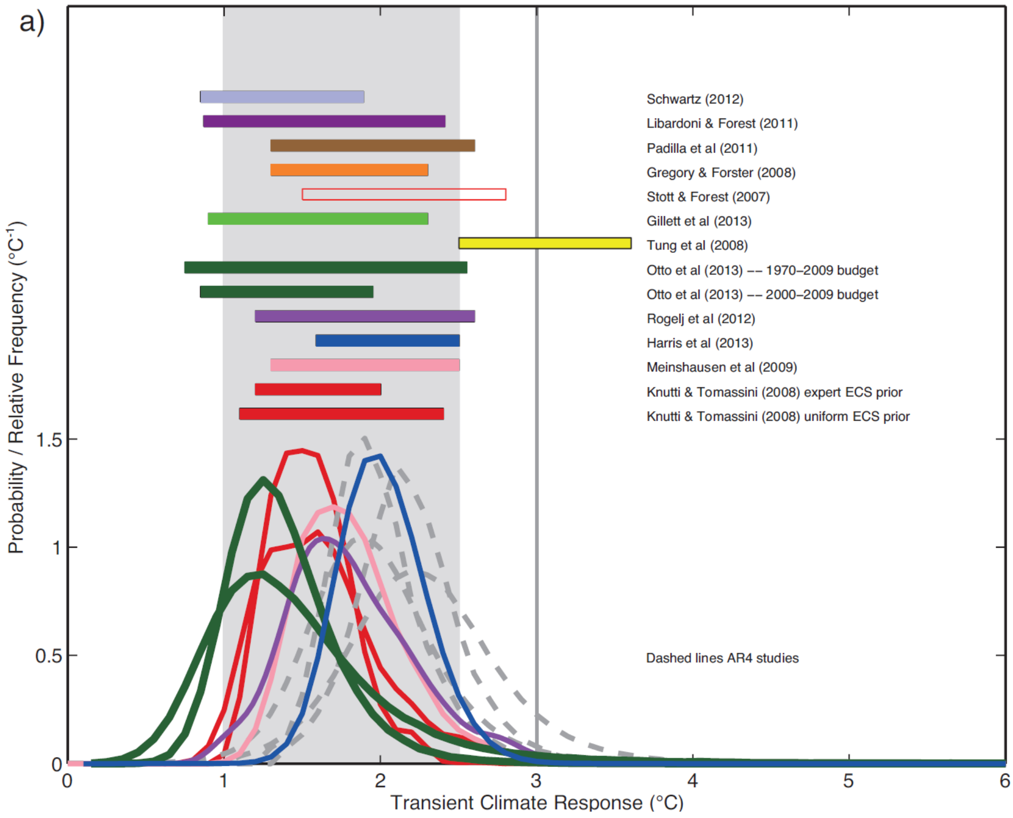

Figure 5: Transient climate response distributions estimated from observational constraints

Reproduced from AR5, Figure 10.20(a). Bars show 5–95% uncertainty ranges for TCR.

estimate for TCR was given in either report. However, a best estimate of 1.3 or 1.4°C for TCR (depending on whether aerosol forcing is scaled to match the satellite observation derived best estimate thereof or not) can be derived from information[86] in the AR5 SPM about changes over 1951–2010, a well-observed period.

---

[86] Dividing the mid-range estimated contributions per Section D.3 of the SPM of greenhouse gases (0.9°C) and other anthropogenic forcings (−0.25°C) to global mean surface temperature over 1951–2010, totalling 0.65°C, by the estimated change in total anthropogenic radiative forcing between 1950 and 2011 of 1.72 W/m^2 per Figure SPM.5, reduced by 0.04 W/m^2 to adjust to 1951–2010, implies a TCR of 1.4°C after multiplying by an $F_{2\times\mathrm{CO}_2}$ of 3.71 W/m^2. When instead basing the estimate on the linear trend increase in observed total warming of 0.64°C over 1951–2010 per Jones et al. (2013) – the study cited in the section to which the SPM refers – (the estimated contribution from internal variability being zero) and the linear trend increase in total forcing per AR5 of 1.75 W/m^2 the implied TCR is also 1.4°C. Scaling the AR5 aerosol forcing estimates to match the mean satellite-observation-derived aerosol forcing estimate would reduce the mean of these two TCR estimates to 1.3°C.

All the good quality observational evidence thus supports a best estimate for TCR of between 1.3 and 1.4°C;[87] taking the mid-point of 1.35°C seems most appropriate. Based on 5–95% ranges for the Gillett et al. (2013), Otto et al. (2013) and Schwartz (2012) studies – faults having been found with all the other estimates shown in Figure 10.20(a) of AR5 – a 'likely' range for TCR of 1–2°C appears suitably conservative.[88]

By contrast, CMIP5 climate model TCRs are on average 35% higher than 1.35°C, at 1.8°C or so, with the TCR for particularly sensitive models substantially higher than that (the UK Met Office HadGEM2-ES model has a TCR of 2.5°C).

Figure 6 compares the best empirical estimate for TCR with the TCR values of the 30 climate models covered in AR5 Table 9.5. There is an evident mismatch between the observational best estimate and the model range. Nevertheless, AR5 states (Box 12.2) that:

> …the ranges of TCR estimated from the observed warming and from AOGCMs agree well, increasing our confidence in the assessment of uncertainties in projections over the 21st century.

How can this be a fair conclusion, when the average model TCR is 35% higher than an observationally-based best estimate of 1.35°C, and almost half the models have TCRs 50% or more above that level? The IPCC obscured this large discrepancy between models and observations by not showing a graph like our Figure 6 and by a misleading statement in the full report.[89]

What will the future bring?

In the SPM, the AR5 report presented projections for global surface temperature increase through to 2100 based on four scenarios for future greenhouse

---

[87] Although not a peer reviewed result, it is worth noting that the well-respected climate scientist Isaac Held argues that TCR is unlikely to exceed 1.8°C, and puts forward a best estimate of 1.4°C. See http://www.gfdl.noaa.gov/blog/isaac-held/2012/04/30/27-estimating-tcr-from-recent-warming

[88] Even the Otto et al. (2013) estimate based on 1970–2009 data, which gives the widest (0.7–2.5°C) of the 5–95% ranges from the three named studies, gives a 17–83% 'likely' range of 1.0–1.9°C.

[89] http://climateaudit.org/2013/12/09/does-the-observational-evidence-in-ar5-support-itsthe-cmip5-models-tcr-ranges/

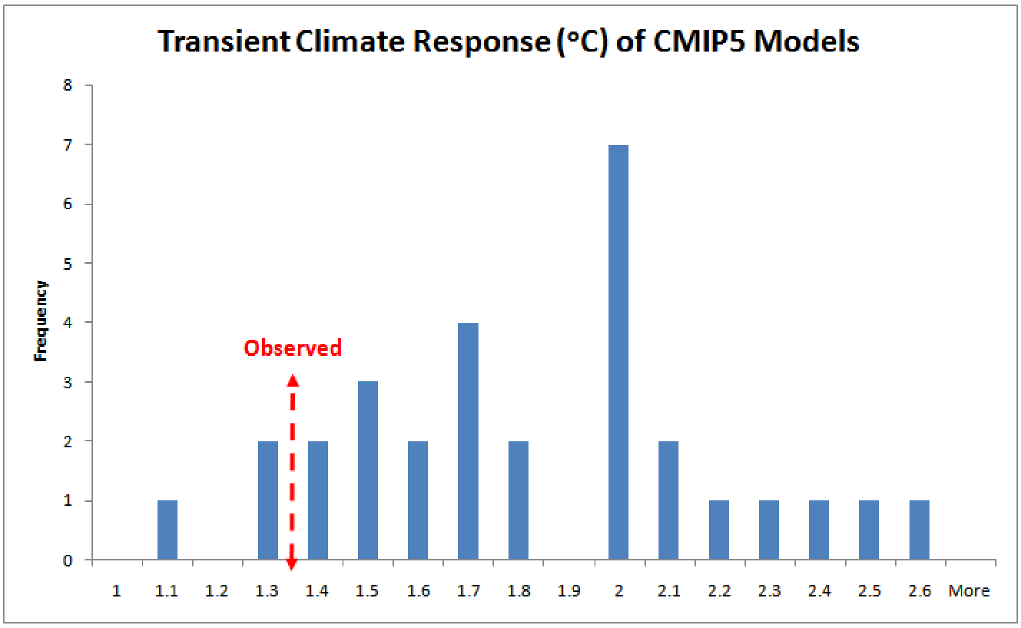

Figure 6: Transient climate response distribution for CMIP5 models

Models per AR5 Table 9.5. The bar heights show how many models in Table 9.5 exhibit each level of TCR.

gas emissions and hence concentrations. These projections are based on simulations by the CMIP5 AOGCMs. Figure 7 below shows (reproduced from Figure SPM7) the projections for two of the scenarios. RCP8.5 is the highest scenario and RCP2.6 is the lowest. Recent increases in greenhouse gas concentrations have been close to those in the middle two scenarios, RCP4.5 and RCP6,[90] although emissions appear to have been increasing at a rate at or above that per the RCP8.5 scenario.

The CMIP5 models estimate warming over the next two decades as a range of 0.48–1.15°C, over all scenarios.[91] In the AR5-WG1 final draft, however, that estimate was reduced by 40% to 0.3–0.7°C, apparently recognising that overall the models were warming unrealistically quickly. Inconsistently, no change was made to the longer-term GCM projections. That results in a jump in projected temperatures between 2016–2035 and 2046–2065.

---

[90] Emissions, and the resulting greenhouse gas concentrations, do not diverge significantly between the RCP4.5 and RCP6 scenarios until after 2050.

[91] 2016–2035 relative to 1986–2005.

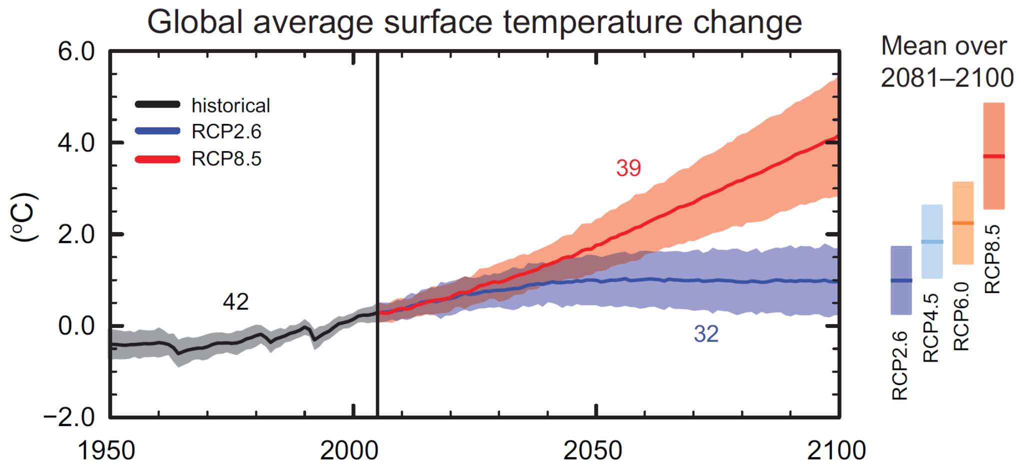

Figure 7: Projected global temperature changes over the rest of the century

Reproduced from AR5, Figure SPM.7. Temperature changes are from the 1986–2005 mean, which was 0.6°C above preindustrial (taken as the 1850–1900 mean global surface temperature). The figures denote the number of models involved.

Observationally-based vs. model projected warming

As we will show, the mean CMIP5 projected warming to 2081–2100 is far above warming projected using the 'best observational' estimate for TCR we derived earlier.[92] In Table 3 we show for each scenario the amount of warming projected in AR5 up to 2081–2100, based on the different scenarios that the IPCC uses, from a baseline of 1850–1900 and also from 2012 (after deducting actual warming from 1850–1900 to 2012). The first two columns show the average warming projected by the CMIP5 climate models. The next two columns show the warming based on the best observational estimate for TCR of 1.35°C. These numbers scale the TCR estimate pro rata to the projected increase in total forcing from 2012 until 2081–2100 on each scenario and then add an allowance for currently unrealised warming from past greenhouse gas increases plus, where relevant, the amount of warming up to 2012. The rightmost column shows the ratio of CMIP5-model to observational-TCR based warming from 2012.

---

[92] The global warming estimates are based on multiplying the TCR estimate of 1.35°C by the change in total forcing on each scenario between 2012 and 2081–2100 per the RCP forcings dataset, and adding 0.15°C for unrealised warming attributable to existing forcing, which as at 2012 was heating the ocean, becoming realised by 2081–2100. These TCR-based projections are consistent with more sophisticated calculations using a 2-box model. Using the mean temperature for the decade ending in 2012 instead of that for 2012 would make no difference.

Table 3: Global warming up to the late twenty-first century

Scenario	**Warming in 2081–2100 based on:**				**Ratio of CMIP5- to TCR-based warming**
	CMIP5 models		**TCR of 1.35°C**		
	°C	°C	°C	°C	
Baseline	1850–1900	2012*	1850–1900*	2012	2012
RCP2.6	1.6	0.8	1.0	0.2	3.4×
RCP4.5	2.4	1.6	1.6	0.8	2.0×
RCP6.0	2.8	2.0	2.0	1.2	1.7×
RCP8.5	4.3	3.5	2.9	2.1	1.7×

*To minimise rounding discrepancies, 0.8°C has been deducted from the CMIP5 global mean surface temperature projected warming from 1850–1900 (taken as representing preindustrial conditions) to obtain warming from 2012, and 0.8°C added to the warming based on TCR from 2012 to obtain warming from 1850–1900. But the unrounded 0.76°C temperature rise from 1850–1900 to 2012 per HadCRUT4 has been used to compute the ratios of CMIP5 model to TCR-based warming.

It can be seen that the climate models greatly overestimate the amount of warming in the future relative to what a best observationally-based estimate of TCR implies. Comparing the two sets of projections of future warming (from 2012 to 2081–2100), and excluding the low RCP2.6 scenario, the model-based projected warming is between 1.7 and 2.0 times higher than the projected warming based on the best observational estimate of TCR. On the RCP6.0 scenario and using the TCR-based method, total warming in 2081–2100 would still be around the international target of 2°C, with a rise of 1.2°C from 2012 rather than the 2°C rise projected by the GCMs.

This exercise reveals a fact that is not evident from AR5: many CMIP5 models simulate faster increases in global surface temperature, particularly in the future, than the model TCR values indicate. While the average model TCR value is 1.8°C or so – 35% higher than our best observational estimate for TCR of 1.35°C, the rise in temperature over the rest of this century projected by the CMIP5 models is much more than 35% higher than that projected on the same scenarios based on a TCR of 1.35°C. Using data on simulated warming over similar periods for all the CMIP5 models analysed in Forster et al. (2013), model average effective TCRs of 2.0°C over the instrumental period, and 2.2°C from the 2000s to the 2090s, can be estimated.[93] Figure 8 shows these TCR values.

---

[93] Effective TCRs over the instrumental period (effective historical TCRs) are estimated for each model as the average of the simulated global surface temperature increases from the start of the simulation (1850 or 1860) to 2001–05 per the historical run and to 2008–12 per the RCP4.5 run, divided by the average increase in total forcing on the RCP4.5 scenario per the RCP dataset over the same periods,

The relationship in many models of effective TCR to the actual model TCR is not stable, and for a majority of models is higher in the twenty-first century.

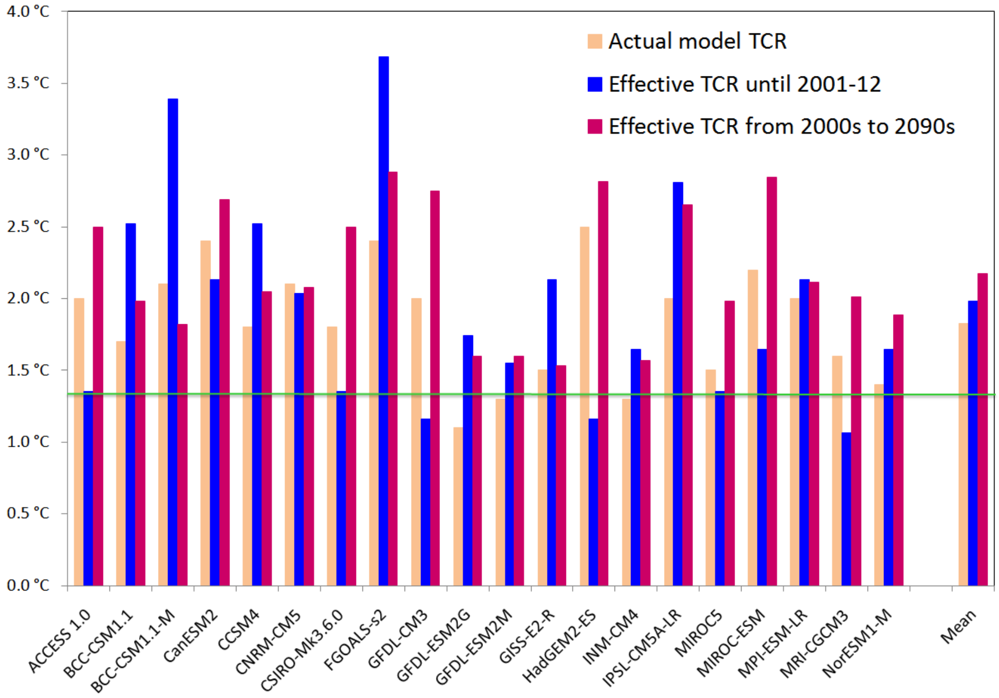

Figure 8: Effective TCRs for CMIP5 models analysed in Forster et al. 2013

The salmon bars show model TCRs, estimated as per the definition of TCR, by increasing the CO_2 concentration in the model by 1% p.a. for 70 years. The blue bars show TCR estimates based on model simulations over the instrumental period, to 2001–12, i.e. 'effective historical' model TCRs. The magenta bars show TCR estimates based on the subsequent change in model-simulated temperatures to 2091–99 on the RCP8.5 scenario, i.e. 'effective future' model TCRs. See footnote 92 for details of their calculation. The multimodel means are shown at the right. The horizontal green line shows the observationally-based TCR best estimate of 1.35°C.

There are probably several reasons for this behaviour. One is that in typical CMIP5 models more warming-in-the-pipeline may well emerge between now

---

and multiplied by an $F_{2\times\mathrm{CO_2}}$ of 3.71 W/m^2 to convert to an effective TCR. The effective TCRs from the 2000s to the 2090s are calculated similarly, based on the subsequent changes in model-simulated temperature to 2091–99 on the RCP8.5 scenario and the corresponding change in mean total forcing on the RCP8.5 scenario per the RCP dataset. Each of the model temperature changes in that future period has been reduced by 0.15°C to allow for the rise that it is estimated would occur in the real climate system by the 2090s in response to past forcing increases even without any future increase in forcing. These two definitions of effective TCR comply reasonably with the generic TCR definition in AR5 Section 10.8.1 provided of the order of 0.15°C warming-in-the-pipeline emerges in the models.

and 2091–2099 than the amount implied by the observationally-based ECS and TCR best estimates. Another is that in some models effective climate sensitivity increases during the first century of simulated warming, reflecting climate state dependency and/or non-linearities in model response. A third is that under the RCP scenarios aerosol emissions are projected almost to halve by 2100 at the same time as greenhouse gas emissions increase strongly. This means that models generating values for aerosol forcing that are higher than per the RCP forcing dataset estimates will simulate a smaller past increase, but a greater future increase, in total forcing than the RCP forcings dataset. Had the RCP forcings used in these calculations been adjusted to be consistent with the best estimates in AR5, which would primarily involve making aerosol forcing less negative,[94] the average excess of estimated model effective future TCRs over effective historical TCRs would have been considerably greater.

Conclusions

Climate science has been under attack in recent years. A major blow to the credibility of the field was Climategate, the hacking of thousands of emails of scientists working at the Climatic Research Unit at the University of East Anglia, among them several lead authors of IPCC reports. The emails showed some of the scientists were trying to keep sceptical studies out of the peer-reviewed literature and IPCC reports and that they were obstructing Freedom of Information requests. Soon afterwards errors were discovered in the AR4 report, of which the melting of the Himalayan glaciers in 2035 was the most visible. All of the errors made climate change 'worse', indicating a bias in the IPCC process.

The climate science community insisted that the errors were all in the Working Group II report, which focuses on the impacts of climate change, and that no errors were found in the Working Group I scientific report. Even the InterAcademy Council, which investigated the IPCC process,[95] concluded that the key findings of the report were still valid: the climate is changing and humans are the cause.[96]

The Working Group I reports contain few outright errors of the magnitude of the one relating to the Himalayan glaciers. Inadequacies in its assessment are

---

[94] The intermittent volcanic forcing would have had to be strengthened, but that would have little effect on the calculation.

[95] http://reviewipcc.interacademycouncil.net/

[96] A curious conclusion as the IAC was not asked to review the science but only the IPCC process and organisation.

more subtle but can also be far more important. In this report we have shown that the AR4 report in 2007 misrepresented an important observational estimate for climate sensitivity, thereby suggesting a higher value for climate sensitivity than the original research indicated and thus making the climate change problem seem 'worse'. Perhaps more importantly, this episode suggests that IPCC authors did not have an adequate grasp of the Bayesian statistical methods used in estimating climate sensitivity.

In the recently released AR5 report the IPCC had the chance to bring policymakers some good news. The highest quality observational evidence indicates climate sensitivity is probably close to the lower bound of the range for climate sensitivity that has been prevalent over the last thirty years, and below the increased lower bound set in the AR4 report. However, as we have shown, the IPCC did not report this news in the clearest possible terms.

IPCC lead authors are bound by the limits of the IPCC process. The IPCC guidelines say that the purpose of the IPCC is 'to assess on a comprehensive, objective, open and transparent basis the scientific, technical and socio-economic information relevant to understanding the scientific basis of risk of human-induced climate change'.[97] 'Comprehensive' means that the authors have to take all the published literature into account, save where it has been superseded or shown to be doubtful. This is what the IPCC authors did in AR5, including climate sensitivity estimates from a large number of published studies (see our Figure 2), including many that we believe are unsatisfactory. They did, however, focus on information provided by recent (instrumental period) observations of changes in, primarily, temperature and also recognised that constraining aerosol forcing and ocean heat uptake was critical to estimating ECS. We agree, but would add the requirement that the methodology be statistically sound. For many studies this was not the case.

In our view, the observational assessments of ECS and TCR in AR5 should accordingly have been informed primarily by those estimates based on observed changes in temperature over the instrumental period that incorporated adequate, observationally based constraints on aerosol forcing and (for ECS estimates) ocean heat uptake, and used sound methodology, particularly as regards statistical methods. That would have meant discounting every one of the high ECS and TCR estimates included in respectively Figures 2 and 5. The studies that qualified would have supported observationally-based best estimates for ECS and TCR substantially below the average values exhibited by

---

[97] http://www.ipcc.ch/organization/organization_history.shtml.

global climate models, and a lower observationally-based range for ECS than the AR5 range.

We have speculated about the reasons for the decision not to give a best estimate for climate sensitivity in AR5. The growing discrepancy between estimates based on models versus those based on observations seems to be at the heart of the matter. It seems likely that the IPCC did not want to put the spotlight on this discrepancy as it would suggest that policy makers should regard projections by the models of future climate change with suspicion.

The IPCC process of being 'comprehensive' allows the authors to stay away from the clear statement that we have made in this report, namely that the best evidence suggests climate sensitivity is close to the reduced, 1.5°C, lower bound. Figure 2 (IPCC Figure 1 in Box 12.2) gives the impression that even just taking the observational studies, many support much higher values of climate sensitivity. We have shown the weakness in these studies. However, if their weaknesses have not been documented in peer reviewed papers, it is difficult for IPCC authors to reject individual studies out of hand. In this case 'bad papers' and those using model-based aerosol forcing estimates helped to obscure the issue, leading to a wider spread of observational estimates of climate sensitivity.

In conclusion, we believe that, due largely to the constraints imposed by the climate-model-orientated IPCC process, the WGI report and the SPM failed to provide an adequate assessment of climate sensitivity – either ECS or TCR – arguably the most important parameters in the climate debate. In particular, it did not draw out the divergence that has emerged between ECS and TCR estimates based on the best observational evidence and those embodied in global climate models.

Acknowledgements

The authors are very grateful for the help and comments they have received from Dr James Annan, Professor Judith Curry, Professor David Henderson, Professor Ross McKitrick and Andrew Montford.

Appendix
Critiques of some observationally-based ECS estimates featured in AR5 Box 12, Figure 1

Libardoni and Forest (2013) – Combination

This model–observation comparison Bayesian study (actually a corrigendum to a study originally published in 2011) uses an informative 'expert' prior distribution for ECS and an inappropriate uniform prior distribution for ocean heat uptake efficiency (the square root of ocean effective diffusivity, K_v). Use of such prior distributions will have biased, most probably upwards, the study's ECS estimate. Using one surface temperature dataset, Libardoni and Forest find ECS to be lower, K_v to be completely unconstrained, and aerosol forcing to be more negative, than the other two datasets are used. Yet with greenhouse gas forcing being offset to a greater extent by negative aerosol cooling and more heat being absorbed by the ocean, energy conservation implies that ECS would need to be significantly higher to match the twentieth-century rise in global temperatures, not lower. Since the Libardoni and Forest results thereby defy conservation of energy, they should be discounted. Although various errors pointed out in Lewis (2013) were addressed in this corrigendum, at least one was incorrectly dealt with, and the unsatisfactory way surface temperature data was used (see Lewis, 2013) was not altered, which may account for these problems.

Lin et al. (2010) – Instrumental

Although this study is dealt with in AR5 alongside studies that involve satellite-measured interannual and interseasonal changes in TOA radiative imbalance, it is really an energy budget study that uses numerical solutions of an energy balance model. The recent TOA imbalance is derived from an outdated AOGCM-derived Earth system heat uptake/TOA radiative imbalance estimate (Hansen et al. 2005) of 0.85 W/m^2, taken as applying over the final decade of the 1885–2005 period used. That heat uptake is twice as high as the best estimate per AR5 over the same decade. Moreover, no allowance is made for heat inflow into the ocean at the start of the 120-year period. The method and model used, in particular the treatment of heat transport to the deep ocean, is difficult to follow and appears non-standard. In view of the greatly excessive system heat uptake estimate used and the questionable methodology, it is difficult to regard the results of this study as constituting a realistic estimate of ECS. The IPCC authors evidently also had doubts about this study's ECS estimate; its range is marked as being incomplete at both high and low ends.

Olson et al. (2012) – Instrumental and combination

This model–observation comparison Bayesian study estimates ECS, ocean effective diffusivity and an aerosol-forcing scaling factor, using only global temperatures and a wide uniform prior on the aerosol-forcing scaling factor. That is an unsatisfactory method. Since greenhouse gas and aerosol forcing histories are extremely closely correlated (negatively), one can obtain a good match to historical global temperatures with a wide range of suitable combinations of ECS and aerosol forcing strength. That problem results in the study's estimated PDF for ECS being almost unconstrained when using uniform prior distributions, which biases its ECS estimate upwards. The use of 0–700-m ocean, as well as surface, temperature changes provides only a very weak constraint on what ECS–aerosol-forcing combinations are feasible. Ozone forcing, which is significantly positive, was omitted: that can be expected to have increased the estimate of ECS substantially. Given all these problems, the Olson et al. instrumental ECS estimate cannot be regarded as realistic.

Olson's PDF and range for ECS shown under combination estimates is dominated by a non-uniform prior distribution for ECS that matches high AR4-era estimates for ECS, including from AOGCMs, as represented in Knutti and Hegerl (2008). Since the study's combination ECS estimate is dominated by an initial distribution based on AR4-era ECS estimates, it should not have been treated in AR5 as if it were an independent observationally-based estimate. The Olson et al. combination estimate for ECS should therefore be disregarded.

Schwartz (2012) – Instrumental

This study derived ECS from changes up to 2009 in observed global surface and 0–700-m ocean layer temperatures, and changes in forcing based on forcing histories used in historical model simulations. Two methods were used. One was zero-intercept regression of temperature change on forcing minus heating rate, fitted to post-1964 data. Whilst this approach appears reasonable in principle, subject to the forcing and OHC history estimates being realistic, the regressions are very noisy. No allowance was made for heat inflow into the ocean in the late nineteenth century (estimated in Gregory et al. 2002, to be non-negligible); that can be expected to have biased upwards its estimate of ECS slightly. For two of the six forcing datasets used, the regressions did not explain any of the variance in the temperature data – their R^2 values were negative. ECS best estimates derived from the other four forcing datasets varied between 1.1°C and 2.6°C.The mean R^2 value for their regressions approached a value of 0.5. The second method derived ECS by combining the results of similar regressions (but without deducting the heating rate from forcing) with an observationally-estimated heat uptake coefficient. These regressions gave significantly higher R^2 values. The second method gave similar results for the four forcing datasets for which the first method provided a valid estimate of ECS, with an overall range (allowing for regression uncertainty) of 1.07–3.0°C. A fifth forcing dataset, which gave a positive R^2 only for the regression in which the heating rate was not deducted, gave an ECS estimate using this method of 4.9±1.2°C. That accounts for the ECS range for this study given in Box 12, Figure 1 of AR5 extending up to 6.1°C. The regression R^2 for this forcing dataset was low (0.29) and the study concluded that the forcing dataset was inconsistent with an energy balance model for which the change in net emitted irradiance at the top of the atmosphere is proportional to the increase in surface temperature. The 3.0–6.1°C segment of the ECS range given for this study in AR5 relates entirely to this one forcing dataset and, in view of the problems with it, should be regarded as carrying significantly less than the one-fifth total probability that would otherwise naturally be assigned to a part of a range that related only to one out of five datasets.

Tomassini et al. (2007) – Instrumental

The Tomassini et al. model–observation comparison study involved a complex subjective Bayesian method. For ECS, a set of priors varying between a uniform prior and a deliberately informative lognormal prior with a mean of 3°C, both restricted to the range 1–10°C, were used. A very inappropriate uniform prior was employed for ocean effective diffusivity (K_v) – the square of ocean heat uptake efficiency. The choices of prior for ECS and K_v will both have biased upwards the estimate of ECS. Although the method used encompasses inverse estimation of aerosol forcing via a scaling factor, only global mean observational temperature data is used, so the inverse estimate arrived at will be unreliable. The very high (negative) correlation between the time evolution of greenhouse gas and aerosol forcings on a global scale makes it impossible robustly to distinguish between different combinations of ECS and aerosol forcing values that each satisfy the energy budget constraint. The posterior distribution for K_v is multiply peaked, which should not be the case. The trace plot of the Markov chain Monte Carlo sample used to estimate the parameters reveals instability not only as to what K_v values are favoured but also as to with what combination of ECS and (indirect) aerosol forcing. In some sections of the plot it is not obvious that the combination of K_v, ECS and aerosol forcing values is consistent with conservation-of-energy constraints. In view of all these issues the ECS estimates from this study should be discounted.

Unlabelled AR4 combination studies

The first unlabelled AR4 study range shown in AR5 Box 12.2, Figure 1 is from Annan and Hargreaves (2006), which is based on a combination of estimates from a last glacial maximum palaeoclimate study and from a study based on the response to volcanic eruptions, using a prior (initial) distribution which peaks at 3°C and has a 2.5–97.5% range of 1–10°C. Since AR5 deprecates ECS estimates based on both these methods and also because the prior distribution used strongly favours high ECS values, no weight should be put on the results. The other unlabelled AR4 range is from Hegerl et al. (2006), which combined its own last-millennium proxy estimate with an instrumental estimate from a modified version of Frame et al. (2005). Problems with these studies, in particular Frame et al. (2005), were described above in the context of the PDFs in Figure 9.20 of AR4. The Aldrin et al. (2012) combination estimate, which likewise uses a last-millennium proxy-based estimate from Hegerl et al. (2006), gives a much lower and better constrained ECS range – showing that the palaeoclimate estimate used has little influence – and is much to be preferred.

Glossary/list of acronyms

AOGCM	Atmosphere–ocean coupled general circulation model
AR4	IPCC fourth assessment report, published in 2007
AR5	IPCC fifth assessment report, published in 2013/2014. Except where the context requires otherwise, references to AR5 are to the finalised version of the AR5 Working Group I report.
Best estimate	This refers to the median estimate, unless otherwise stated.
CMIP3 models	Generation of AOGCMs used to provide simulation runs (CMIP3 runs) for AR4.
CMIP5 models	Generation of AOGCMs used to provide simulation runs (CMIP5 runs) for AR5.
ECS	Equilibrium climate sensitivity, the change in the annual mean global surface temperature, once the deep ocean has come into equilibrium, following a doubling of the atmospheric carbon dioxide concentration (or a change in the overall mixture of greenhouse gases that causes the same change in forcing). It does not reflect adjustment by components of the climate system with even slower timescales (e.g. ice sheets or vegetation).
Effective climate sensitivity	An estimate of equilibrium climate sensitivity that is evaluated from non-equilibrium conditions. The two terms are treated in this report as synonymous, and both referred to as ECS, as is generally the case in AR5.
Forcing	See RF.
GCM	General circulation model, a mathematical model of the general circulation of a planetary atmosphere(or sometimes ocean). Also referred to, along with AOGCMs, as global climate models.
IPCC	Intergovernmental Panel on Climate Change
LGM	Last glacial maximum, the time of maximum extent of ice sheets during the last glaciation, approximately 21,000 years ago

Mean, median, mode	Different types of central estimates for data. One obtains the mean by adding up all the data values and then dividing by the number of data points. The median is the middle value in the data set, with equal numbers of lower and of higher values. The mode is the value that appears most often. For continuous data, or an uncertain parameter, having a probability density (PDF) rather than discrete values, the mean is derived by integrating the value of the data or uncertain parameter over the PDF; the median is the 50th percentile of the probability distribution, where the probabilities of the data or uncertain parameter having higher or lower values are equal (the value at which the area under the PDF with higher values equals the area under it with lower values); and the mode is the value at which the PDF peak is located. Distributions for climate sensitivity are often substantially asymmetrical (skewed). For such distributions the median is generally accepted as being a better central or best estimate than the mode or the mean.
Ocean heat uptake efficiency	A measure of how rapidly heat is absorbed by the ocean below the relatively shallow (averaging of the order of 100 m) mixed layer. For a diffusive ocean model the relevant efficiency measure is the square root of ocean effective vertical diffusivity.
OHC	Ocean heat content, the heat stored in the ocean.
PDF	Probability density function, a function that describes the relative likelihood for a variable to take on a given value. The function integrates to unity over the entire range that the variable may possibly take, and its integral (the area under the PDF) over any sub-range indicates the probability that the actual value of the variable lies within that sub-range.

RF	Radiative forcing (often just forcing), the change in TOA net radiative balance caused by a change in CO_2 concentration or in any other external driver of climate change. It is expressed in units of watts per square metre (W/m^2 or Wm^{-2}). The term is used in this report, as in AR5, to refer to effective radiative forcing (ERF), a concept that includes the effects of rapid non-surface temperature climate system adjustments to the change in radiative forcing.
SPM	Summary for Policymakers (pertaining to AR5 WGI, unless the context requires otherwise).
TCR	Transient climate response, defined as the change in the global mean surface temperature, averaged over a 20-year period, centred at the time of atmospheric carbon dioxide doubling, in a climate model simulation in which CO_2 increases at 1% per annum compound, which takes 70 years. The value of TCR can be derived using a different rate of increase in CO_2 over 70 years, by scaling the change in global temperature inversely. TCR can be more easily estimated than ECS, and is more relevant to projections of warming – although not sea level rise – over the rest of this century.
TOA	Top-of-atmosphere.
Troposphere	The lowest part of the atmosphere, from the surface to about 10 km in altitude at mid-latitudes, where clouds and weather phenomena occur. Above the troposphere lies the stratosphere.
WGI	IPCC Working Group One

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.

Andrews, D. and M. Allen, 2008. Diagnosis of climate models in terms of transient climate response and feedback response time, *Atmos. Sci. Lett.*, 9: 7–12.

Andronova, N.G. and M.E. Schlesinger, 2001. Objective estimation of the probability density function for climate sensitivity. *J. Geophys. Res.*, 106: 22605–22611.

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

Annan, J.D. and J.C. Hargreaves, 2011. On the generation and interpretation of probabilistic estimates of climate sensitivity. *Clim. Change.*,104, 423-436.

Domingues, C.M., J.A.Church, N.J. White, P.J. Gleckler, S.E. Wijffels, P.M. Barker and J.R. Dunn, 2008. Improved estimates of upper-ocean warming and multidecadal sea-level rise. *Nature*, 453: 1090–3.

Forest, C.E., P.H. Stone, A.P. Sokolov, M.R. Allen and M.D. Webster, 2002. Quantifying uncertainties in climate system properties with the use of recent climate observations. *Science*, 295: 113–117

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

Forest, C.E., P.H. Stone, and A.P. Sokolov, 2008. Constraining climate model parameters from observed 20th century changes. *Tellus A.*

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.

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

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

Fyfe, J.C., N.P. Gillett, and F.W. Zwiers, 2013. Overestimated global warming over the past 20 years. *Nature Clim. Ch.*, 3.9: 767–769.

Gillett, N.P., V.K. Arora, D. Matthews, P.A. Stott, and M.R. Allen, 2013. Constraining the ratio of global warming to cumulative CO_2 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.

Hansen, J., Nazarenko, L., Ruedy, R., M. Sato, J. Willis, A. Del Genio, D Koch, A. Lacis, K. Lo, S. Menon, T. Novakov, J. Perlwitz, G. Russell, G. Schmidt and N. Tausnev, 2005. Earth's energy imbalance: Confirmation and implications. *Science*, 308: 1431–1435.

Harris, G.R., D.M.H. Sexton, B.B.B. Booth, M. Collins, J.M. Murphy and M.J. Webb, 2006. Frequency distributions of transient regional climate change from perturbed physics ensembles of general circulation model simulations. *Clim. Dynam.*, 27: 357–375.

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.*, 40: 2937–2972.

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.

Jaeger, C.C., and J. Jaeger, 2011. Three views of two degrees. *Reg. Env. Change*, 11, S15–S26.

Jewson, S., 2013. Methods for the inclusion of parameter uncertainty in weather and climate forecasts. Presentation at the European Meteorological Society 11th European Conference on the applications of meteorology. Available at: http://presentations.copernicus.org/EMS2013-151_presentation.pdf.

Jones, G.S., P.A. Stott and N. Christidis, 2013. Attribution of observed historical near surface temperature variations to anthropogenic and natural causes using CMIP5 simulations. *J. Geophys. Res. Atmos*, doi:10.1002/jgrd.50239.

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 Geosc.*, 1: 735–743.

Levitus, S., J. Antonov, T. Boyer and C Stephens, 2000. Warming of the world ocean, *Science*, 287: 5641.2225–2229.

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., L. Chambers, P. Stackhouse Jr., B. Wielicki, Y. Hu, P. Minnis, N. Loeb, W. Sun, G. Potter, Q. Min, G. Schuster and T.-F. Fan, 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, 2009. On the determination of climate feedbacks from ERBE data. *Geophys. Res. Lett.* 36, L16705.

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.

Loeb, N.G., J.M. Lyman, G.C. Johnson, R.P. Allan, D.R. Doelling, T. Wong, B.J. Soden and G.L. Stephens, 2012. Observed changes in top-of-the-atmosphere radiation and upper-ocean heating consistent within uncertainty. *Nature Geoscience.*, 5: 110–113.

Masters, T., 2013. Observational estimate of climate sensitivity from changes in the rate of ocean heat uptake and comparison to CMIP5 models. *Clim. Dyn.*, DOI 10.1007/s00382–013–1770–4

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 Geosci.*, 6: 415–416.

Paleosens Members, 2012. Making sense of palaeoclimate sensitivity. *Nature*, 491: 683–691.

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.

Schlesinger, M., H. Kheshgi, J. Smith, F. de la Chesnaye, J.M. Reilly, T. Wilson, and Ch. Kolstad, 2007. *Human-induced Climate Change: An Interdisciplinary Assessment*. Cambridge University Press.

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.

Soden, B.J. and I.M. Held, 2006. An assessment of climate feedbacks in coupled ocean-atmosphere models. *J. Clim.*, 19: 3354–3360.

Tol, R.S.J., 2007. Europe's long-term climate target: A critical evaluation. *Energy Policy*, 35: 424–432.

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.

Van der Sluijs, J.P., J.C.M. van Eijndhoven, B. Wynne, and S. Shackley, 1998. Anchoring devices in science for policy: The case of consensus around climate sensitivity. *Soc. Studies Sci.*, 28; 2: 291–323.

VonderHaar, T.H., J.L. Bytheway, and J. M. Forsythe, 2012. Weather and climate analyses using improved global water vapor observations. *Geophys. Res. Lett.*, 39: L15802, doi:10.1029/2012GL052094.

The Global Warming Policy Foundation is an all-party and non-party think tank and a registered educational charity which, while open-minded on the contested science of global warming, is deeply concerned about the costs and other implications of many of the policies currently being advocated.

Our main focus is to analyse global warming policies and their economic and other implications. Our aim is to provide the most robust and reliable economic analysis and advice.

Above all we seek to inform the media, politicians and the public, in a newsworthy way, on the subject in general and on the misinformation to which they are all too frequently being subjected at the present time.

The key to the success of the GWPF is the trust and credibility that we have earned in the eyes of a growing number of policy makers, journalists and the interested public.

The GWPF is funded overwhelmingly by voluntary donations from a number of private individuals and charitable trusts. In order to make clear its complete independence, it does not accept gifts from either energy companies or anyone with a significant interest in an energy company.

Published by the Global Warming Policy Foundation

For further information about the GWPF or a print copy of this report contact:

The Global Warming Policy Foundation
10 Upper Bank Street, London E14 5NB
T 020 7006 5827
M 07553 361717
www.thegwpf.org

Registered in England, no 6962749
Registered with the Charity Commission, no 1131448