Comments on the recent Met Office report

"The recent pause in global warming (3): What are the implications for projections of future warming?"

Nicholas Lewis[1]

Independent climate scientist

September 2013

[1] email: nhlewis@btinternet.com

Introduction and Summary

These comments constitute a response to erroneous statements and misrepresentations made in a report published by the Met Office in July 2013: "The recent pause in global warming (3): What are the implications for projections of future warming?" (the Report).

A significant part of the Met Office Report concerns what it refers to as a recent comprehensive study, being the 2013 Nature Geoscience paper "Energy budget constraints on climate response" (Otto et al). The Report contains several misrepresentations of the findings of the Otto et al study as well as a number of other misleading or erroneous statements.

The author team of Otto et al includes fourteen lead or coordinating lead authors of the forthcoming IPCC scientific report (AR5 WG1 report). The author of this memorandum, Nicholas Lewis, whilst being one of the authors of the Otto et al 2013 study, writes in a personal capacity, not as a representative of its author team.

The Met Office Report discusses in some detail estimates of transient climate response (TCR) and equilibrium/effective climate sensitivity (ECS). TCR, a measure of the rise in global surface temperature at the end of a 70 year period over which atmospheric CO_2-equivalent concentrations grow at 1% per annum (and hence double), is closely linked to projections of human-caused warming several decades into the future. ECS measures the eventual surface temperature increase from such a doubling once ocean temperatures have fully adjusted[1]. ECS largely determines TCR[2].

The Met Office Report considers whether climate model estimates of TCR and ECS need to be revised in the light of recent observational evidence, in particular the relatively slow increase in global surface temperature over the last 15 to 20 years. It concludes that, broadly, such estimates – and hence climate model projections of future warming – do not need to be reduced.

One of the Met Office Report's main conclusions is that "the upper ranges of TCR and ECS derived from extended observational records … are broadly consistent with the upper range from the latest generation of comprehensive climate models" (CMIP5 models). This is contradicted by Otto et al, which stated "Our results match those of other observation-based studies and suggest that the TCRs of some of the models in the CMIP5 ensemble with the strongest climate response to increases in atmospheric CO_2 levels may be inconsistent with recent observations". Barely half CMIP5 models analysed[3] have TCRs below the 95% upper bound for TCR of 2.0°C given in Otto et al, using observational data for the latest decade.

These issues are of particular relevance to the Met Office. Both the TCR and ECS of its flagship HadGEM2-ES model[4], used for policy advice, are very near the top of the range for

---

[1] ECS approximates to what TCR would have been in the absence of absorption of heat by the ocean, which slows down the rise in temperature.

[2] in conjunction with ocean heat absorption characteristics

[3] Forster et al 2013

[4] http://www.metoffice.gov.uk/climate-change/policy-relevant/advance

CMIP5 models[3]. Its TCR exceeds the 95% bound derived from CMIP5 models other than HadGEM2-ES[5]. The HadGEM2-ES model's TCR also exceeds the 95% bound derived from previous generation (CMIP3) models. The Met Office HadGEM2-ES model's TCR of 2.5°C is not only well above the upper 95% bound of 2.0°C given in Otto et al[6], but also above the 2.3°C bound given in Gillett et al 2013 – the only other study cited in the Met Office Report that derived TCR from observational records. Indeed, the HadGEM2-ES TCR is nearly double the Otto et al best estimate for TCR of 1.3°C[6]. As for ECS, the HadGEM2-ES model's ECS of 4.6°C lies well beyond the upper 95% bounds given not only by Otto et al[6] but also by Aldrin et al 2012 (3.5°C[7]) and Lewis 2013 (3.0°C[8]). HadGEM2's ECS also exceeds the 4.5°C top of the IPCC's 'likely' range, and the 95% bounds both from CMIP3 models and from CMIP5 models excluding HadGEM2-ES[5].

The TCR of the Met Office's previous generation climate model, HadCM3, is at the 95% upper bound for TCR of 2.0°C given in Otto et al. A perturbed physics ensemble study based on adjusting the parameters in HadCM3, now set out in two major published papers (Sexton et al 2012 and Harris et al 2013), represents the techniques and climate model on which the official UK Climate Projections[9] produced by the Met Office were based. HadCM3's high TCR might not matter if, as the Report claims, "uncertainty in the response of the climate system to CO_2 forcing is comprehensively sampled" in the study. However, it is not. Despite thorough attempts to make it do so by varying its parameters, HadCM3 is unable to sample the region where, according to several recent observational studies, key characteristics of the real climate system are most probably located. It is therefore unsurprising that incorporating observational data barely alters the Harris et al 2013 prior central estimate for TCR.

Misrepresentations relating to transient climate response (TCR)

The Met Office Report refers to three methods of estimating TCR: from simulations made with climate models, from observations, and by combining climate model and observationally-derived values. It makes the contentious claim that none of these methods can be said to be superior to the others. In science, it is standard to test the validity of theoretical models by comparing their predictions to observational data. Accordingly, it seems possible to say that estimates derived purely from simulations by climate models, without combination with observationally-derived values, are likely to be inferior to those from the other two methods.

Figure 1 and Table 1

Figure 1 in the Report, which gives the Otto et al TCR 5–95% range as 0.7 to 2.5°C, most likely value 1.4°C, based on 1970–2009 data, is objectionable on a number of grounds:

---

[5] Forster et al 2013, recalculating the multi-model means and standard deviations with HadGEM2-ES excluded

[6] primary, arguably most reliable, estimate based on data for the latest decade (2000–09)

[7] Aldrin et al 2012 main results

[8] Lewis 2013 preferred main results after incorporation of uncertainty in non-aerosol forcings and surface temperature observations

[9] UKCP09: Murphy et al 2009. *UK Climate Projections Science Report: Climate change projections*

1. Figure 1 does not use the Otto et al primary TCR best estimate of 1.3°C and 5–95% range of 0.9 to 2.0°C, based on data for the decade 2000–09. Although caution is required in interpreting results for any short period, arguably – as stated in Otto et al – the estimate based on the most recent decade's data is the most reliable since it has the strongest forcing and is much less affected by the 1991 eruption of Mount Pinatubo. Accordingly, showing in Figure 1 only the (wider) TCR estimated range based on 1970–2009 data for Otto et al for comparison with other estimates is misleading.

2. Classifying the Harris et al 2013 HadCM3-perturbed-physics-ensemble derived TCR estimates as based on model and observations gives a misleading impression of the relative influence of those two factors. Although highly sophisticated, the study would more appropriately be classified as primarily model based. As HadCM3's parameters are perturbed, the resulting changes in ECS and aerosol forcing are closely linked. When significantly lower values for ECS – as suggested by recent observational studies – are obtained, HadCM3's aerosol forcing takes on highly negative values[10]. The observational data strongly contraindicate aerosol forcing being highly negative, so parameter combinations resulting in significantly reduced model ECS levels (and thus highly negative aerosol forcing) are heavily down-weighted. As a result, whatever the actual level of ECS, HadCM3-derived ECS estimates are bound to be high. See Box 1. The HadCM3-derived TCR estimates are very largely determined by its ECS estimates, so the same is true for TCR. Since the Harris et al 2013 results largely reflect the particular characteristics of the HadCM3 model, and are at variance with results from several recent fully observationally-constrained studies, no reliance should be placed upon them.

3. The central TCR estimate of 1.6°C given for Gillett et al 2013 is a mean, and is not comparable to the median estimates quoted for Otto et al and Harris et al 2013. Since TCR and ECS probability distributions are usually skewed, the median – the 50th percentile of the distribution – is a much preferable central estimate for TCR and ECS than the mean (the probability-weighted average value). Gillett et al 2013 used a standard detection-and-attribution method, giving regression coefficients which can then be multiplied by the model TCRs to produce observationally-based TCR estimates. The median TCR estimate based on individual model regression coefficients was 1.46°C[11].

4. The multi-model CMIP3 and CMIP5 central TCR estimates given of 1.8°C are also means rather than medians. The median TCR for the set of CMIP3 models referenced is lower, at 1.6°C. (The mean–median difference is negligible for the set of CMIP5 models.)

[10] The close relationship between aerosol forcing and ECS levels in HadCM3 is shown in Figure S2 of Harris et al 2013 and reflected in the estimated relationship between them given in its Equation (S9). The finding in Harris et al 2013 that use of an alternative basis, in which aerosol forcing is treated as independent of ECS, has little effect on its results is misleading. That finding is of little relevance, since it is only during the second stage in the generation of the results that aerosol forcing is treated as being independent of ECS, by which point the ECS estimates are largely determined. Moreover, the IPCC AR4 aerosol forcing distribution used under the alternative basis is considerably more strongly negative than is supported by current best estimates.

[11] Gillett et al used a single multi-model based regression coefficient, which would give a median 0.04°C higher

Box 1: Harris et al 2013 – the Met Office's HadCM3-based PPE study

Harris et al 2013 is based on an ensemble of many simulations of transient and equilibrium warming by the HadCM3 model using differing values for key parameters, so producing a range of different model characteristics. Using a complex subjective Bayesian statistical methodology, the prior probability distribution for these perturbed physics ensemble (PPE) simulations is sampled and weighted according to how well each matches large-scale metrics of recent mean climate observations, producing a mean-climate-weighted distribution. The weighted ECS estimates are then used to drive a simpler climate model (SCM) in which further uncertainties, including in aerosol forcing[12], are sampled. The estimates from the ensemble of HadCM3 simulations are then scaled by global mean temperatures predicted by the SCM. Comparisons of the resulting ensemble of scaled simulations with historical changes in observed surface temperature are then used to reweight the mean-climate-weighted distribution, producing final 'posterior' probability distributions for global surface temperatures, ECS and TCR. (A more complex scaling method is used for other variables.) Considerable allowances for uncertainties are made. However, the final probability distribution for ECS (and hence TCR) is largely determined during the first stage: the weighting by changes in observed temperatures has little effect on what are by then fairly tight ranges for ECS and TCR. The methodology depends on the key assumption that the spread of plausible PPE projections is consistent with an adequate range of possible values for climate system properties. In HadCM3 both ECS and aerosol forcing are largely controlled by the same parameters and so have a close relationship[13]. At parameter settings that reduce HadCM3's standard ECS of 3.3°C to 2°C or below, its aerosol forcing becomes very negative. Observations of mean climate render such ECS-aerosol forcing combinations implausible. The HadCM3 PPE is unable to simulate a climate system in which both climate sensitivity and aerosol forcing are relatively modest, the region where estimates by several recent observational studies are centred. Figure B.1 illustrates this point.

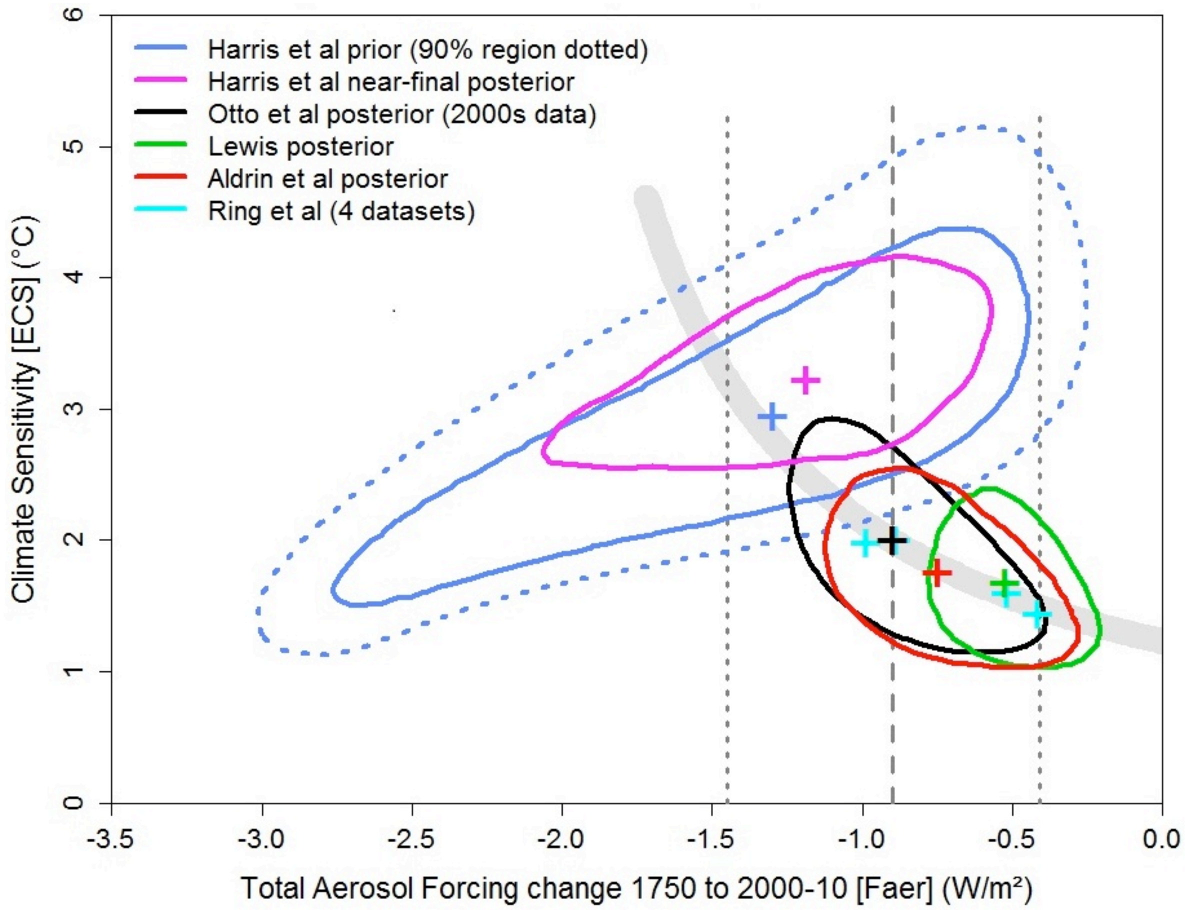

Figure B.1: ***Contours enclosing 'likely' (66%) probability regions in the Climate Sensitivity – Aerosol Forcing plane***. Showing Harris et al 2013 estimated prior and mean-climate-weighted distributions; final posterior distributions for Otto et al 2103, Lewis 2013[14] and Aldrin 2012. The joint medians for each distribution are shown by crosses. For Ring et al 2012 the crosses show the best estimates resulting from use of four different surface temperature datasets. The thick grey line indicates approximately how best estimates of ECS may be expected to vary with aerosol forcing given the latest observational estimates of changes in global surface temperature, non-aerosol forcings and the Earth's energy imbalance from 1860–79 to 1990–2009. A best estimate of recent total aerosol forcing is shown by the dashed grey vertical line, with the related 'likely' range bounded by dotted grey lines.

[12] Aerosols are thought to have a cooling influence, quantified by the extent to which aerosol forcing is negative.

[13] See Harris et al 2013 Supplementary Information equation (S9), used to estimate the joint prior distribution shown in Figure B.1

[14] preferred main results as per note 8. Aerosol forcing rescaled to 2000-10 and adjusted to exclude black-carbon-on-snow forcing

A guide to interpreting Figure B.1

- The graph shows the relationship between probabilistic estimates of climate sensitivity and aerosol forcing for recent studies for which the necessary information is available. The reason for showing this relationship is that uncertainty regarding aerosol forcing accounts for much of the variation in best estimates of climate sensitivity.
- The relevant solid contour line for each study encloses the highest probability density area comprising 66% of the total probability. The intersection of the 50th percentiles of the probability distributions for each of the two variables – their 2D joint median – is shown by a cross.
- The grey curve shows the approximate dependence of best estimates of climate sensitivity on best estimates of aerosol forcing: that which would be expected if studies reflected the best observational estimates of changes over the last 150 years in global climate system variables other than aerosol forcing. The thickness of the curve indicates that there is a range of uncertainty as to best estimates, but does not purport to quantify it.
- The vertical dashed grey line shows a current best estimate for aerosol forcing, with the dotted grey lines showing the related 17–83% ('likely', in IPCC terminology) uncertainty range.
- For all studies, but less so for Harris et al 2013, the best estimates per their posterior distributions (the final observationally-weighted distributions except for Harris et al 2013, where the distribution post weighting by mean climate observations is shown) are consistent with the expected relationship between best estimates of aerosol forcing and of climate sensitivity: their crosses lie on or close to the thick grey curve.
- For the observationally-based studies, their contour regions are elongated in the direction of the grey curve, indicating that the largest uncertainty is in aerosol forcing, as expected. The bulk of all their 'likely' regions lie within the estimated 'likely' range for aerosol forcing shown by the dotted grey lines. Their 2D medians all lie very close to the grey curve. (Otto et al used a satellite-based best estimate of aerosol forcing, equal to the best estimate per Figure B.1, rather than deriving its own estimate: its cross lies exactly on the grey dashed line.)
- There is a large overlap between the 'likely' regions (for Ring et al 2012, the four best estimates based on different surface temperature datasets) from all the observationally-based studies.
- In all the observationally-based studies, there was no link between the *prior* distributions (the initial assumptions before weighting by observational data) for aerosol forcing and for climate sensitivity. Nor were any values for those variables that were remotely compatible with the observations ruled out initially, although aerosol forcing values close to zero were given a very low prior weighting in the Aldrin et al 2012 study.
- In Harris et al 2013, there is a very strong link between the *prior* distributions for aerosol forcing and for climate sensitivity, shown by the regions enclosed by the blue contour lines. The solid and dotted blue contours enclose respectively 66% and 90% of the estimated prior probability weighting. *The region initially sampled by Harris et al 2013 excludes low climate sensitivity, modest aerosol forcing combinations – the region where the observationally-based studies' best estimates lie.* There is not even a significant overlap between the initial Harris et al 2013 'likely' region and the posterior 'likely' regions produced by the observationally-based studies.
- At its core (highest probability density point) the Harris et al 2013 prior distribution has a weighting over one hundred times as high as it does at the core of the Otto et al. posterior distribution. The down-weighting by the Harris et al 2013 prior is even greater at the cores of the posteriors of the other observationally-based studies.
- Weighting the Harris et al 2013 prior distribution to reflect recent mean climate observations heavily down-weights strongly negative aerosol forcing values. This shift downwards in the strength of aerosol forcing would be expected, from the shape of the grey curve, to result in a reduction of estimated climate sensitivity. But here, perversely, estimated climate sensitivity increases. That reflects the strong, opposing, relationship between aerosol forcing and climate sensitivity that exists in HadCM3 as its parameters are varied, shown by its prior.
- The Harris et al 2013 climate sensitivity estimate post weighting by recent mean climate observations is fairly well constrained, and so is little changed by the further weighting by observations of historical surface temperature changes – even when at that stage an ECS-independent, aerosol forcing distribution model is tried out as an alternative. The TCR estimates are largely a function of the ECS estimates and so also change little.
- The conclusion is that, despite the great sophistication of Harris et al 2013, its weighting by both recent mean climate and historical temperature change observations and the careful effort put into sampling uncertainties, **the study is fundamentally unsatisfactory because it effectively rules out from the start the possibility that both aerosol forcing and climate sensitivity are modest. And that is the combination that recent observations are supporting, according to a number of peer-reviewed studies**.

Several of the TCR estimates given for Otto et al in Table 1 are wrong. The central TCR estimate based on the 2000s observational period is given as 1.4°C. It was actually 1.3°C. Several of the other TCR percentiles given are also wrong[15].

A revised version of Figure 1 in the Met Office Report showing the impact of the revisions discussed above, and a better idea of the distribution of probability, is shown in Figure 1.

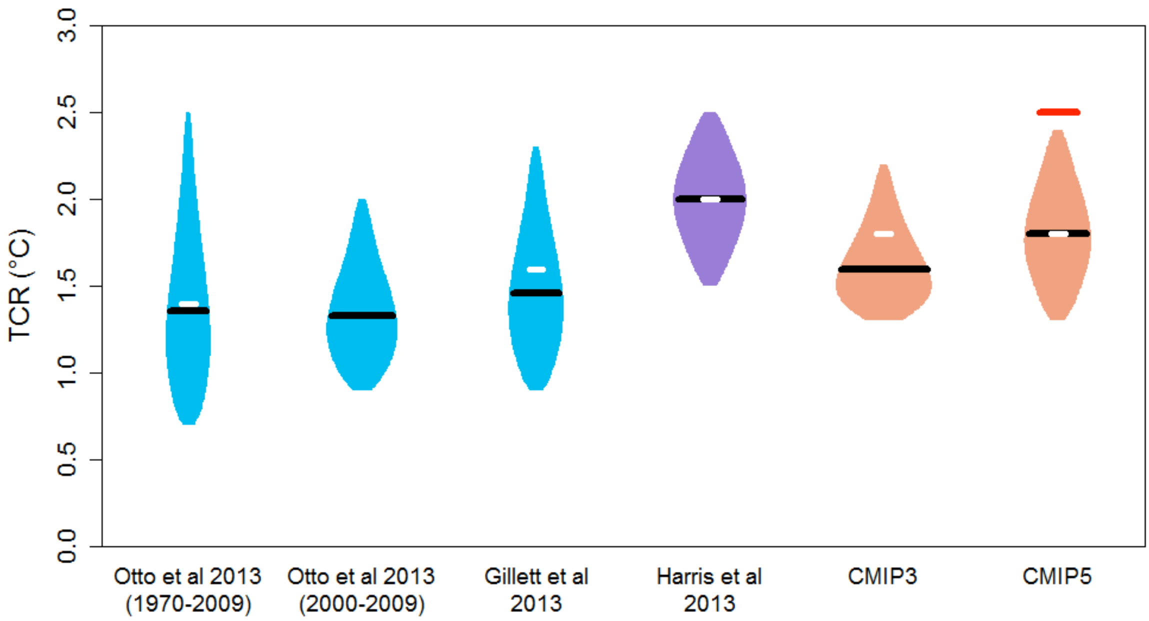

Figure 1: ***Estimates of TCR from the same sources as in Figure 1 of the Met Office Report, with estimates based on 2000-2009 data added for Otto et al.***

The flasks span the 5–95 % uncertainty ranges for the estimates from each source; the black horizontal lines mark their 50th percentiles (medians). The width of the flask shows probability density, in each case for the shifted lognormal distribution corresponding to the applicable 5th, 50th and 95th percentiles. Since flasks from all sources therefore have equal area, probability is proportional to area between as well as within flasks. Short white horizontal bars show the central estimate for each source given in Figure 1 of the Met Office Report. The red bar in the CMIP5 column shows the TCR of the Met Office HadGEM2-ES model. The colours denote type of source, with blue showing estimates based on observations, purple showing estimates based primarily on model but partly on observations, and salmon for model-only estimates.

Physical considerations

The Met Office Report states that "To reach the very low values [for TCR] quoted in Otto et al. (2013) would require negative feedbacks to be acting quite strongly to counteract the well understood physics of greenhouse gas radiative forcing, water vapour feedback and surface

---

[15] The 5th percentiles for the 1970s and 1980s observational periods are respectively 0.1°C and 0.6°C, not 0.3°C and 0.7°C. And the 95th percentile for the 1970s observational period is 3.1°C not 3.0°C. Otto et al TCR estimate percentiles for the 1970s and 1980s observational periods can be computed accurately from the data given in its Supplementary Information.

albedo feedback". That is misleading. The strong negative lapse rate feedback is very closely linked to the water vapour feedback (they are sometimes combined into a single feedback) and has a similar level of understanding. Therefore, it should be included along with greenhouse gas radiative forcing, water vapour feedback and surface albedo feedback.

A multi-model study of feedbacks, Soden and Held 2006, showed a median ECS for the model ensemble of 1.8°C after combined water vapour/lapse rate and surface albedo feedbacks. The median 1.3°C and 1.4°C estimates for TCR in Otto et al both correspond to ECS estimates above 1.8°C, so are consistent with the Soden and Held 2006 findings without requiring any additional negative feedbacks to be acting. Moreover, real-world water vapour feedback may not be as strong as in typical climate model simulations. Although the basic physics of these feedbacks may be well understood, there remains substantial uncertainty as to their magnitudes. Furthermore, cloud feedbacks are highly uncertain.

Estimated warming at 2100

The Met Office Report miscalculates estimated warming at 2100 under the RCP8.5 scenario for all the studies shown, mainly by using inconsistent bases in the calculation. Figure 2 of the Report also contains an additional misrepresentation of Otto et al's results. Both issues can be illustrated using the values for Otto et al. The 5–95% range given in Figure 2 of the Report for Otto et al, of 1.7–6.2°C with a median value of 3.5°C, is evidently based on multiplying the study's TCR estimates based on 1970–2009 data by 8.5 Wm^{-2} and dividing by 3.44 Wm^{-2}, the CMIP5 mean F_{2x}[16] level in Forster et al 2013, as used in Otto et al. The RCP8.5 scenario was so named because the resulting forcing approaches 8.5 Wm^{-2} in 2100. However, that was based on a different, higher, F_{2x} basis. In fact, the RCP8.5 indicative forcings dataset (Meinshausen et al, 2011) gives a level of 8.34 Wm^{-2} in 2100, not 8.5 Wm^{-2}, and it is based on a canonical F_{2x} estimate of 3.71 Wm^{-2}, not 3.44 Wm^{-2}. So the RCP8.5 forcing in 2100 should be divided by 3.71 Wm^{-2}, not 3.44 Wm^{-2}, in order to estimate warming in 2100. Substituting that divisor in the calculation, and taking the indicative RCP8.5 mean forcing[17] over 2095-2105 to even out the solar cycle, reduces the 3.5°C median 2100 warming estimate from Otto et al using 1970–2009 data to 3.1°C.

Moreover, the Met Office Report does not reveal that the warming at 2100 shown for Otto et al is based on its TCR estimates using data for 1970–2009 rather than its primary, arguably more reliable, TCR estimates based on data for the 2000s. Using those estimates instead, the median 2100 warming estimate from Otto et al under the RCP8.5 scenario becomes 2.9°C, with a 5–95% range of 2.0–4.4°C[18]. Both the median estimate and the top of the 95% range are substantially lower than the figures of respectively 3.5°C and 6.2°C given in Figure 2 of the Report. Note that the statement in the caption thereto that medians are marked is incorrect for 3 out of 5 sources; it is means that are marked, except for Otto et al and Harris et al 2013.

[16] F_{2x} is the radiative forcing (warming influence) caused by a doubling of atmospheric CO_2 concentrations.

[17] of 8.25 Wm^{-2}. RCP8.5 data is available at http://www.pik-potsdam.de/~mmalte/rcps/index.htm#Download

[18] An alternative calculation using the RCP8.5 CMIP5 multimodel mean forcing in 2095 from Forster et al 2013 and using (correctly, in this case) 3.44 Wm^{-2} as the divisor produces a slightly lower range of 1.9–4.3°C, with a reduced median of 2.8°C.

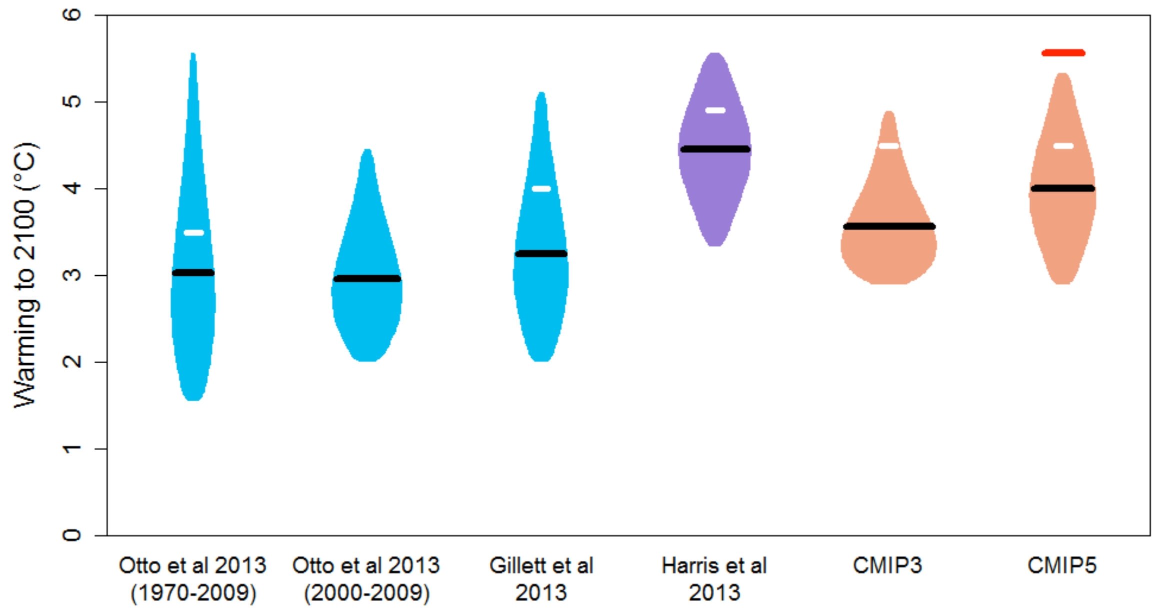

Figure 2: ***Estimated warming at 2100 relative to pre-industrial levels under the RCP8.5 scenario.*** Values are calculated as in Figure 2 of the Met Office Report but on a corrected basis and using the TCR values from Figure 1 above rather than from Figure 1 of the Report. The flasks span the 5–95 % uncertainty ranges for the estimates from each source; the black horizontal lines mark their 50th percentiles (medians). The width of the flask shows probability density, in each case for the shifted lognormal distribution corresponding to the applicable 5th, 50th and 95th percentiles. Since flasks from all sources therefore have equal area, probability is proportional to area between as well as within flasks. Short white horizontal bars show the central estimate for each source given in Figure 2 of the Report. The red bar in the CMIP5 column shows the Estimated warming at 2100 for the Met Office HadGEM2-ES model. The colours denote type of source, with blue showing estimates based on observations, purple showing estimates based primarily on model but partly on observations, and salmon for model-only estimates.

The use of inconsistent F_{2x} bases and overstatement of RCP8.5 forcing at 2100 does not just affect the Otto et al study. All the estimated 2100 warming figures given for Gillett et al 2013, Harris et al 2013, CMIP3 and CMIP5 similarly need to be reduced by 10% in order to correct these errors. A revised version of Figure 2 in the Met Office Report showing the impact of these corrections is shown in Figure 2. As noted in the Report, RCP8.5 is the highest of all the RCP scenarios, which should be borne in mind when considering the projected Warming to 2100 levels in Figure 2.

The uncertainty ranges shown both in the Met Office Report Figure 2 and Figure 2 here are in any case exaggerated, since 25–30% of the decadal-mean radiative forcing projected under RCP8.5 for 2100 relative to pre-industrial levels has already occurred. Moreover, according to observational records, by 2012 the global mean temperature had increased by about 0.8°C. That is in line, scaling pro rata, with 25–30% of the medians for the Otto et al and Gillett et al projected increases to 2100. But the non-observational studies' central warming projections need to be adjusted down to reflect their overestimates of the temperature increase to date.

Misrepresentations relating to equilibrium climate sensitivity (ECS)

Estimation of equilibrium climate sensitivity is dealt with in Section 4 of the Met Office Report. It repeats the misleading claim that "Positive feedbacks in the physical climate system, the largest of which is the water-vapour feedback, increase this number [ECS] to over 2°C", and compounds this distortion by stating that "the fundamental physics of climate sensitivity, involving black body radiation and water vapour feedbacks … alone give a climate sensitivity of at least 2.0°C". As already pointed out, these claims ignore the negative lapse-rate feedback, which is intimately linked to the water-vapour feedback. After including lapse-rate feedback as well as water-vapour feedback, all the models analysed in Soden and Held 2006 had a climate sensitivity of below 2.0°C.

The Met Office Report also claims, concerning the Otto et al ECS estimate based on data for the 2000s, that "As for TCR the observations of most recent decade are not representative of the full observational record and so not expected to be representative of the longer term future." That claim is not valid. Decadal-scale internal variability of global temperature was allowed for in Otto et al. Moreover, the slower increase in surface temperatures during the 2000s was, as one would expect, associated with a higher estimated rate of ocean heat uptake, which increases the ECS estimate. The Otto et al ECS estimates based on data for each of the 1980s, the 1990s and the 2000s, and for 1970-2009, measured in each case relative to estimates for 1860–79, were all very similar. Indeed, the Otto et al ECS estimate based on 2000–09 data was actually the highest of those four ECS estimates, as a result of the top-of-the-range ocean heat uptake estimate used.

Figure 5 of the Met Office Report only gives ECS 5-95% bounds for Otto et al and CMIP3 and CMIP5 model ensembles. It is surprising that no other recent peer-reviewed observationally-based estimated ranges for ECS were shown. Examples are Aldrin et al 2012, Lewis 2013 and Masters 2013.

It is questionable whether valid estimates of probabilistic ranges can be constructed from an ensemble of simulations by different climate models. But even assuming so, Figure 5 of the Report indicates it does not show, as stated, the 5-95% ECS range for CMIP5 models given by Forster et al 2013. That range is 1.9–4.5°C, not 2.1–4.6°C. Moreover, the CMIP3 and CMIP5 central estimates are means, whereas those for Otto et al are medians.

A revised version of Figure 5 in the Met Office Report that does include several recent observationally-based studies, and corrects the CMIP5 ECS 5-95% range, is shown in Figure 3. All central estimates are medians to provide consistency. The range from palaeoclimate estimates (1.2–5.2°C) has been omitted since it is not on a comparable basis to the remaining ranges and palaeoclimate estimates involve great uncertainties.

It is evident from Figure 3 that a variety of recent observationally-based studies all produce median estimates for ECS that are very substantially below the CMIP3 and CMIP5 model-based estimates, indicating that ECS in the more sensitive CMIP3 and CMIP5 models, at least, is likely to be out of line with reality.

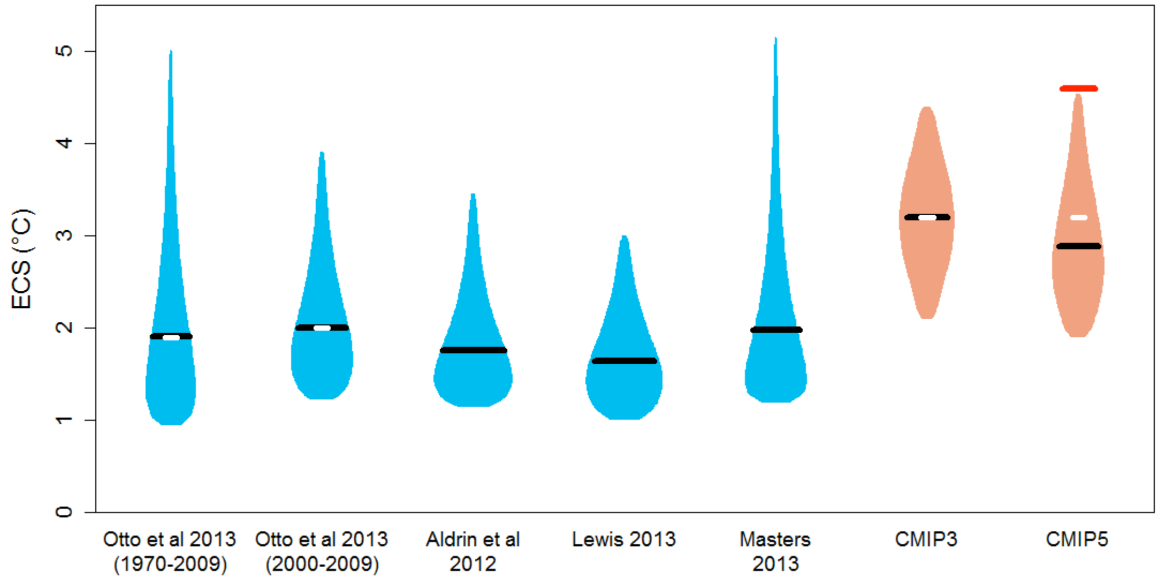

Figure 3: ***Estimates of ECS from the same sources as in Figure 5 of the Met Office Report, with those from Aldrin et al 2012, Lewis 2013[8] and Masters 2013 added and the Palaeo range omitted***. The flasks span the 5–95 % uncertainty ranges for the estimates from each source; the black horizontal lines mark their 50th percentiles (medians). The width of the flask shows probability density, in each case for the shifted lognormal distribution corresponding to the applicable 5th, 50th and 95th percentiles. Since flasks from all sources therefore have equal area, probability is proportional to area between as well as within flasks. Short white horizontal bars show the central estimate for each source given in Figure 5 of the Report. The red bar in the CMIP5 column shows the ECS of the Met Office HadGEM2-ES model. The colours denote type of source, with blue showing estimates based on observations and salmon for model-only estimates.

Conclusions

The Met Office Report contained a number of misrepresentations of Otto et al 2013, misleading statements and outright mistakes. Correcting these and providing a fairer analysis reveals that, although some of the recent observationally-based estimates of TCR and ECS considered have upper bounds to their uncertainty ranges that are comparable to those for model-based estimates, they all give lower best estimates than model-based best estimates. Observationally-based median estimates for TCR and ECS are often comparable to the bottom of model-based uncertainty ranges.

The statement in the Met Office Report that "the upper ranges of TCR and ECS derived from extended observational records … are broadly consistent with the upper range from the latest generation of comprehensive climate models" has been shown to be questionable. The TCR and ECS characteristics of the Met Office flagship HadGEM2-ES model are of particular interest in this regard. None of the 95% bounds for TCR estimates from the observationally-based sources cited in the Report exceed the TCR of HadGEM2-ES. And HadGEM2-ES has an ECS that exceeds not only the 95% bound from Otto et al but also that from two other recent observationally-based studies. Moreover, both the TCR and the ECS of HadGEM2-ES exceed the 95% bounds derived not only from CMIP3 models but also from CMIP5 models other than HadGEM2-ES.

It has also been shown that estimates based on perturbing parameters in the Met Office HadCM3 model, as exemplified by Harris et al 2013, do not sample combinations of ECS and aerosol forcing located in the region most supported by several recent observationally-based studies, and so cannot be regarded as properly reflecting observational evidence. That is of concern since these same techniques, using the HadCM3 model, represent the basis on which the official UK Climate Projections were constructed.

References

Aldrin, M., Marit Holden, Peter Guttorp, Ragnhild Bieltvedt Skeie, Gunnar Myhre and Terje Koren 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.

DelSole, T., M. Tippett, and J. Shukla, 2011: A significant component of unforced multidecadal variability in the recent acceleration of global warming. Journal of Climate, 24 (3), 909–926.

Forster, P. M. et al., 2013: Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models. J. Geophys. Res., doi:10.1002/jgrd.50174 .

Gillett, N.P., V. K. Arora, G. M. Flato, J. F. Scinocca, K. von Salzen, 2012: Improved constraints on 21st-century warming derived using 160 years of temperature observations. Geophys. Res. Lett, 39, L01704, doi:10.1029/2011GL050226.

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

Harris, Glen R., David M. H. Sexton, Ben B. B. Booth, Mat Collins and James M. Murphy, 2013: Probabilistic projections of transient climate change. Clim. Dyn. 40, 2937-2972. DOI 10.1007/s00382-012-1647-y.

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, doi:10.1002/jgrd.50239.

Lewis, Nicholas, 2013: An objective Bayesian, improved approach for applying optimal fingerprint techniques to estimate climate sensitivity. Journal of Climate, DOI: 10.1175/JCLI-D-00473.1.

Masters, Troy, 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.

Meinshausen, M., S. Smith et al., 2011: The RCP greenhouse gas concentrations and their extension from 1765 to 2500. Climate Change, Special RCP Issue.

Murphy, James M., David M. H. Sexton et al., 2009: UK Climate Projections Science Report: Climate change projections. Met Office Hadley Centre. Available at http://ukclimateprojections.dafra.gov.uk

Otto, Alexander, Friederike E. L. Otto, Olivier Boucher, John Church, Gabi Hegerl, Piers M. Forster, Nathan P. Gillett, Jonathan Gregory, Gregory C. Johnson, Reto Knutti, Nicholas Lewis, Ulrike Lohmann, Jochem Marotzke, Gunnar Myhre, Drew Shindell, Bjorn Stevens, and Myles R. Allen, 2013: Energy budget constraints on climate response. Nature Geosci., doi:10.1038/ngeo1836.

Ring, M.J., D. Lindner, E.F. Cross and M. E. Schlesinger, 2012: Causes of the Global Warming Observed since the 19th Century. Atmospheric and Climate Sciences, 2, 401-415.

Sexton, David M. H., James M. Murphy, Mat Collins and Mark J. Webb, 2012: Multivariate probabilistic projections using imperfect climate models part I: outline of methodology. Clim. Dyn 38, 2513-2542. DOI 10.1007/s00382-011-1208-9.

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