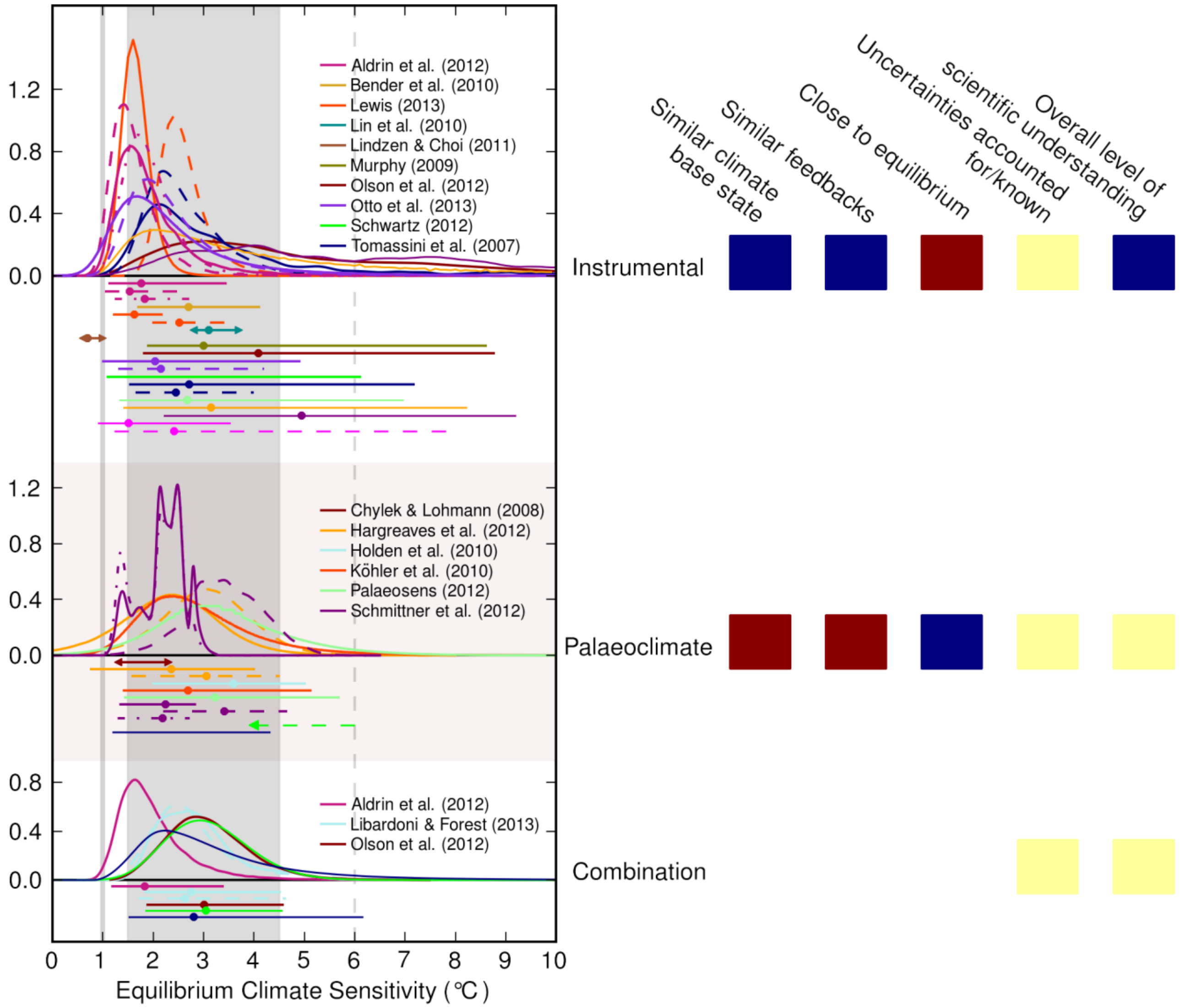
Brief critiques of Instrumental and Combination observationally-based ECS estimates included in AR5 WG1 Figure 10.20.b) that measure surface warming over an extended period.

This document provides critical assessments of probabilistic estimates for equilibrium climate sensitivity (ECS) included in Figure 10.20.b) of the accepted version of the IPCC fifth assessment Working Group 1 report (AR5 WG1), published on 30 September 2013. Figure 10.20.b) is reproduced below.



Reproduction of Figure 10.20.b) from the accepted version of AR5 WG1: **Examples of distributions for ECS estimated from observational constraints.** The overall assessed 1.5–4.5°C 'likely' range is the solid grey band. Selected old estimates used in AR4 (no labels, thin lines) are compared with new estimates available since then (labelled, thicker lines), with 5–95% uncertainty ranges shown by horizontal lines (circles: median values) and distributions (probability densities) shown where available. Ranges that are assessed as being incomplete are marked by arrows. Where available, results are shown using several different prior distributions. The boxes on the right hand side indicate limitations and strengths of each line of evidence: blue indicates high, pale yellow medium, and dark red low, confidence, dependent on how well the evidence is understood, how many studies there are and how well they agree, and how low the uncertainties are and what confidence there is that they are accounted for relatively completely.

All the new AR5 Instrumental and Combination estimates that involved measuring surface warming over an extended period are assessed; typically these studies also used measurements of sub-surface ocean temperatures. As AR5 says, estimating ECS generally relies on the paradigm of a comparison of observed change with results from a physically based climate model, given uncertainty in the model, data, radiative forcing, and due to internal variability. The climate model involved is normally simpler than an atmosphere-ocean general circulation model (AOGCM) and has calibrated variable parameters that control ECS and other key climate system properties separately. Since the model is used to perform simulations over wide ranges of each such property, these are referred to as model-multiple-simulations – observation comparison studies. The goodness of fit between observations and simulations at each parameter value combination indicates how likely it is to correspond to the actual climate system. The resulting likelihood function is usually combined, using Bayes theorem, with a prior distribution for each parameter being estimated, producing a estimated joint PDF for the parameters, from which a PDF and uncertainty range for ECS is obtained by integrating out the other parameters.

The equilibrium to which ECS refers to is generally taken to be an equilibrium involving the ocean-atmosphere system. Estimates of ECS based on temperature change during the instrumental period can only sample atmospheric feedbacks that occur with presently evolving climate change. In that sense they estimate the 'effective climate sensitivity', but in some cases what they estimate is calibrated to ECS as derived from GCM or AOGCM simulations. The ECS of a coupled AOGCM is normally estimated from transient simulations, and so is based on measuring effective climate sensitivity displayed by simulations using the standard AOGCM parameter settings. In this document, no distinction is drawn between estimates of effective and equilibrium climate sensitivity.

A recent AOGCM-based study, Armour et al (2013), claims that regional feedbacks can cause effective climate sensitivity to vary with time, and to be significantly lower than ECS even over periods of several hundred years. However, Williams et al (2008) concluded that after allowance was made for comparatively fast (one year to two decade) forcing adjustments, there was little evidence for time-varying effective climate sensitivity in AOGCMs. They also found that ECS estimated using slab ocean versions of AOGCMs, run to equilibrium, could be significantly inaccurate measures of the ECS of the full ocean model AOGCMs. Armour et al (2013) compared effective climate sensitivity with the ECS of the slab-ocean version of the CCSM4 AOGCM used in their study, which (per AR5 WG1 Table 9.5) is over 10% higher than the ECS of the fully-coupled CCSM4 model. If they had compared the time evolution of effective climate sensitivity with the reported ECS of the fully-coupled CCSM4 model, the two would have become very close within 100–150 years. Moreover, the latitudinal feedback pattern in CCSM4 appears different from that in most other models (see Zelinka and Hartman, 2012), which might account for their results. Dr Armour accepts[1] that the feedback pattern in CCSM4 may be somewhat of an outlier.

[1] http://rankexploits.com/musings/2013/observation-vs-model-bringing-heavy-armour-into-the-war/#more-21626

AR5 WG1 concluded (sections 10.8.2.4 and 12.5.3) that estimates based on past climate states very different from today may differ from the climate sensitivity measuring the climate feedbacks of the Earth system today, and that in any case uncertainties in such estimates were large. Accordingly, estimates in the Palaeoclimate section of Figure 10.20.b) – all of which involved such climate states – are not assessed. AR5 WG1 also concluded (sections 10.8.2.2, 10.8.2.3 and 12.5.3) that measurement and methodological uncertainties in estimates of the feedback parameter and ECS from short term variations in the satellite period preclude strong constraints on ECS; and that the constraint on ECS from the response to volcanic forcing was fairly weak and, being based on a short-term non greenhouse gas forcing, unreliable. Estimates using those methods (respectively Lindzen & Choi (2011) and Murphy (2009); and Bender (2010)), are therefore not considered here.

The old AR4 Instrumental estimates are not dealt with in detail here. Most of the ranges and medians given for them are significantly wrong (see Table 9.3 in AR4 WG1 for the correct ranges). And all the AR4 Instrumental estimates other than that represented by the solid pink line are so badly flawed, by use of inappropriate statistical methods and/or poor data, as to be of little relevance. The dark blue AR4 Combination estimate is from Hegerl et al (2006) and combined its own last-millennium proxy based estimate with a flawed, biased high, instrumental estimate that dominated this Combination estimate. The dominance of the instrumental estimate in the Hegerl et al Combination estimate is evident from the very different range given by the Aldrin et al (2012) Combination estimate, which also used the Hegerl et al (2006) last-millennium estimate. The green AR4 combination estimate is from Annan and Hargreaves (2006). It is based on a combination of estimates from a Last Glacial Maximum paleoclimate study and from a study based on the response to volcanic eruptions – both methods that AR5 expresses reservations about – and uses a prior distribution which peaks at 3°C.

Aldrin et al (2012) – Instrumental and Combination

This is an impressively thorough and very well documented Bayesian study based on observations encompassing almost all the instrumental period. It uses ocean heat content as well as hemispheric surface temperature measurements. This observational data is compared with simulations by a relatively simple but hemispherically-resolving model. The comparison gives rise to a quite well-constrained inverse estimate of aerosol forcing, which is essential for constraining ECS. The Aldrin et al (2012) ranges make allowance for natural internal variability and other sources of uncertainty, including in nonaerosol forcings. Although Aldrin et al (2012) used a subjective Bayesian statistical method, examination suggests that only the priors used for aerosol forcing and ECS appear questionable. The study's inverse estimate of aerosol forcing was biased negatively by the prior used, relative to what the observational likelihood implied, which causes an upwards bias to ECS estimation. However, the bias was moderate except the when the prior was made more strongly negative to allow for a possible cloud lifetime effect

Several PDFs are shown for Aldrin et al (2012) in Figure 10.20.b). The solid line uses a uniform in ECS prior, with dash-dots using the same basis but with data updated to 2010. The dashed line PDF, constrained to lower levels, uses a uniform in 1/ECS prior. The range and median derived using a non-Bayesian profile likelihood method on the estimated joint ECS-aerosol forcing distribution are closely in line with those based on the uniform in 1/ECS prior, which strongly suggests that prior is more appropriate than the uniform in ECS prior.

The Combination estimate shown for Aldrin et al (2012) uses the ECS PDF estimated by the last millennium paleoclimate study Hegerl et al (2006) as its prior for ECS.

Lewis (2013) - Instrumental

This study used simulations by the intermediate complexity MIT 2D model. As the writer was the study's sole author, no detailed assessment of it will be given. Two points should, however, be made.

All the estimated PDFs for ECS in Figure 10.20.b) reflect a Bayesian probabilistic approach. Lewis (2013) was the only Bayesian study that explicitly adopted an objective approach and computed a noninformative prior distribution. (Otto et al, 2013, used a sampling method and hence implicitly used an objective approach with a noninformative prior.) Given the weak observational data and highly non-linear relationships between the data and both ECS and measures of ocean heat uptake efficiency, subjective Bayesian methods (typically employing either wide uniform priors or 'expert' priors) are unlikely to provide estimates that correctly reflect what the data error distributions imply.

The dashed line Lewis (2013) PDF was derived using the Forest et al (2006) data and diagnostics, so as to be comparable with the results of that study. However, as stated in Lewis (2013), on that basis the best estimates of ECS and the other climate system parameters involved produce greatly excessive simulated warming over the 140 year period simulated. Contrariwise, the observed warming is matched when the revised diagnostics (involving longer data and improved design of the surface temperature diagnostic) that give rise to the solid line PDF in Figure 10.20.b) are used. Therefore the solid line PDF, which is based on the revised diagnostics, is much to be preferred to the dashed line PDF. However, the model simulations used in Forest et al (2006), and hence both PDFs shown in Figure 10.20.b), ignored uncertainties in nonaerosol forcings. That study (and hence Lewis, 2013) also ignored uncertainty in surface temperature measurements. As stated in Lewis (2013), incorporation of such uncertainties is estimated, when the revised diagnostics are used, to increase the ECS 5-95% range to 1.0–3.0°C but to leave the median unchanged at 1.6°C.

Libardoni and Forest (2013) – Combination

This Bayesian study, like Lewis (2013), compares observations with multiple simulations by the MIT 2D model at many different parameter settings. It is actually a corrigendum to a study originally published in 2011. The study 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 estimation. Using one surface temperature dataset, Libardoni and Forest find ECS to be lower, K_v to be completely unconstrained above, and aerosol forcing to be more negative, than when using the other two datasets. 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 20th century rise in global temperatures, not lower. Since the Libardoni & Forest results thereby defy basic physical principles, they must 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 rate 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 a little confusing and appears substantially non-standard. In particular, heat uptake by the deep ocean is assumed to be proportional to the rate of change of surface temperature, exactly as for the ocean mixed layer. In view of the greatly excessive system heat uptake estimate used and the very questionable model of heat uptake by the deep ocean, 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 low and high ends.

Olson et al (2012) – Instrumental and Combination

This model-multiple-simulations – observation comparison Bayesian study estimates ECS, ocean effective diffusivity and an aerosol forcing scaling factor, using only global temperatures and with 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 above when using uniform prior distributions, which bias ECS estimation upwards. The use of 0-700m ocean, as well as surface, temperature changes provides only a very weak constraint as to 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.

Otto et al (2013)

This is an energy budget study and so should provide robust estimates. The ECS estimates given in Figure 10.20.b) are based on changes, between 1860–1879 and either 2000–2009 or 1970–2009, in observed global surface temperature, in the Earth's climate system heat content (using the same estimates as in AR5 WG1) and in mean total forcing as simulated by the sample of CMIP5 models analysed in Forster et al (2013). The multi-model mean forcing changes were adjusted to reflect an assessment that estimates based on satellite observations indicate recent aerosol forcing to be some 0.3 W/m^2 less negative than in CMIP5 models. Allowing for the different F_{2xCO2} values (the estimated forcing resulting from a doubling of atmospheric CO_2 concentration) involved, the resulting change in mean forcing between 1860–1879 and 2000–2009 used in Otto et al is closely in line with what the AR5 best estimate of the forcing change between the same periods would be, if adjusted to reflect the mean of the recent satellite-observation derived estimates used to inform estimation of total aerosol forcing in AR5. Since mean total forcing was higher over 2000–2009 than over 1970–2009, and was also less affected by volcanic activity, the ECS estimate based on 2000–2009 data is arguably more reliable than that based on 1970–2009 data.

Note: the writer was a co-author of Otto et al (2013).

Schwartz (2012) – Instrumental

This study derived ECS from changes up to 2009 in observed global surface and 0–700m ocean layer temperatures, and changes in forcing based on forcing histories used in historical model simulations. It gives a composite range based on sampling from various forcing datasets and the estimate and uncertainty band from each, rather than a standard 5–95% range. 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, the regressions are 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 estimation 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), possibly indicating that those two forcing histories are unrealistic. 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 was approaching 0.5. The second method of estimating ECS derived it from 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 explains why the ECS range for this study given in Figure 10.2.b) of AR5 extends up to 6.1°C. The without-deduction-of-heating-rate regression R^2 for this forcing dataset (0.29) was much lower than for any of the other four datasets with positive R^2 values (mean: 0.66), 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 much 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

This model-multiple-simulations – observation comparison study used 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 estimation 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 very 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.

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, AtmosSci Let, 9, 7-12

Annan, J. D., and J. C. Hargreaves, 2006: Using multiple observationally-based constraints to estimate climate sensitivity. Geophysical Research Letters, 33, L06704.

Armour, K. C., C. M. Bitz and G. R. Roe, 2013: Time-varying climate sensitivity from regional feedbacks. Journal of Climate, 26, 4518–4534.

Bender, F. A. M., A. M. L. Ekman, and H. Rodhe, 2010: Response to the eruption of Mount Pinatubo in relation to climate sensitivity in the CMIP3 models. Climate Dynamics, 35, 875-886.

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

Forster, P. M., 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. Journal of Geophysical Research, 118, 1139–1150.

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

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

Lewis, N., 2013: An objective Bayesian, improved approach for applying optimal fingerprint techniques to estimate climate sensitivity. Journal of Climate, 26, 7414-7429.

Libardoni, A. G., and C. E. Forest, 2011: Sensitivity of distributions of climate system properties to the surface temperature dataset. Geophysical Research Letters, 38, L22705.

Libardoni, A. G., and C. E. Forest, 2013: Correction to "Sensitivity of distributions of climate system properties to the surface temperature dataset". Geophysical Research Letters, doi:10.1002/grl.50480.

Lin, B., et al., 2010a: Estimations of climate sensitivity based on top-of-atmosphere radiation imbalance. Atmospheric Chemistry and Physics, 10, 1923-1930.

Lindzen, R. S., and Y. S. Choi, 2011: On the observational determination of climate sensitivity and its implications. Asia-Pacific Journal of Atmospheric Sciences, 47, 377-390.

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. Journal of Geophysical Research Atmospheres, 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. Journal of Geophysical Research Atmospheres, 117, D04103.

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

Schwartz, S. E., 2012: Determination of Earth's transient and equilibrium climate sensitivities from observations over the twentieth century: Strong dependence on assumed forcing. Surveys in Geophysics, 33 745-777.

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. Journal of Climate, 20, 1239-1254.

Williams, K. D., W. J. Ingram and J. M. Gregory, 2008:. Time Variation of Effective Climate Sensitivity in GCMs. Journal of Climate, 21, 5076–5090.

Zelinka, M. D. and D. L. Hartmann, 2012: Climate Feedbacks and Their Implications for Poleward Energy Flux changes in a warming climate. Journal of Climate, 25, 608–624.