

RESEARCH LETTER

10.1002/2013GL058532

Key Points:

- We evaluate climate sensitivity estimates in a perfect model approach
- Climate sensitivity mode varies by 0.8°C (20%) from natural variability
- As the data constrain gets strong, the model structure becomes limiting

Supporting Information:

- Readme
- Text S1

Correspondence to:

M. Huber,
markus.huber@env.ethz.ch

Citation:

Huber, M., U. Beyerle, and R. Knutti (2014), Estimating climate sensitivity and future temperature in the presence of natural climate variability, *Geophys. Res. Lett.*, 41, 2086–2092, doi:10.1002/2013GL058532.

Received 31 OCT 2013

Accepted 7 FEB 2014

Accepted article online 12 FEB 2014

Published online 28 MAR 2014

Estimating climate sensitivity and future temperature in the presence of natural climate variability

Markus Huber¹, Urs Beyerle¹, and Reto Knutti¹¹Institute for Atmospheric and Climate Science, Universitaetstrasse 16, ETH Zurich, Zurich, Switzerland

Abstract We use an initial condition ensemble of an Earth System Model as multiple realizations of the climate system to evaluate estimates of climate sensitivity and future temperature change derived with a climate model of reduced complexity under “perfect” conditions. In our setup, the mean and most likely estimate of equilibrium climate sensitivity vary by about 0.4–0.8°C ($\pm 1\sigma$) due to internal variability. Estimates of the transient climate response vary much less; however, the effect of the spread and bias in the transient response on future temperature projections increases with lead time. Future temperature projections are shown to be more robust for central ranges (i.e., likely range) than for single percentiles. The estimates presented here strongly depend on a delicate balance between a particular realization of the climate system, the emerging constraints on the estimates as well as on the signal, and the decreasing uncertainties in ocean heat uptake observations.

1. Introduction

After more than 40 years, equilibrium climate sensitivity (ECS) is still today’s most widely used metric to characterize the magnitude of the Earth’s temperature response to changes in radiative forcing. Based on different lines of evidence, the Intergovernmental Panel on Climate Change (IPCC) concluded in its recent Fifth Assessment Report (AR5) that the likely range (central >66%) is 1.5–4.5°C [Intergovernmental Panel on Climate Change, 2013], featuring a slightly lower limit compared to the previous Assessment Report (AR4). The upper temperature limit did not change, and the likely range for ECS has been rather robust in the past decades [Knutti and Hegerl, 2008]. Yet the quest for a tighter uncertainty range of equilibrium climate sensitivity recently gained new momentum when studies used the updated observed surface and ocean warming together with energy budget equations [Otto *et al.*, 2013] or reduced complexity climate models [Aldrin *et al.*, 2012; Lewis, 2013] and found values near the lower end of the previously long-standing likely range to be more plausible. In contrast, the current generation of fully coupled climate models of the Coupled Model Intercomparison Project 5 (CMIP5) as well as recent estimates from the paleorecord show still a mean estimate of about 3.2°C ($\pm 1.3^\circ\text{C}$, 90% uncertainty) [Forster *et al.*, 2013] and a range of 2.2–4.8°C [Rohling *et al.*, 2012], respectively. The reason for the different modes for ECS (i.e., median) across different studies, methodologies, and models is still unclear.

Here we focus on probabilistic methods with reduced complexity climate models which infer climate system properties such as ECS from the observed changes in surface temperature and ocean heat content by fitting an ensemble of climate model simulations to the observed record [e.g., Aldrin *et al.*, 2012; Forest *et al.*, 2002; Huber and Knutti, 2012; Knutti *et al.*, 2002; Lewis, 2013]. Both surface and upper ocean temperatures that are used as observational constraints in these methods have increased more slowly in the last decade or two compared to earlier periods, with the period since 1998 often termed the hiatus period. At the same time, the global top-of-atmosphere energy balance still shows a negative imbalance [Loeb *et al.*, 2012; Stephens *et al.*, 2012]. Recent modeling studies suggest that internal variability, including the El Niño–Southern Oscillation and the Interdecadal Pacific Oscillation play a key role in this hiatus, resulting in more heat uptake and transport to the deep ocean [Guemas *et al.*, 2013; Kosaka and Xie, 2013; Meehl *et al.*, 2011, 2013a].

However, a validation of these probabilistic estimates is neither feasible nor is their sensitivity to unforced climate variability known, since we are inevitably bound to use the single realization of the real-world climate system as constraint, with climate sensitivity being unknown. To overcome this issue, we use a 20-member initial condition ensemble of an Earth System Model as multiple realizations of the climate system to evaluate such climate sensitivity estimates under “perfect” conditions and to quantify the effect of internal variability.

The setup used here extends a previous study using a similar approach by *Olson et al.* [2013] in that it includes all historic forcings, the uncertainties associated with individual forcing agents as well as the application of both real-world observations and simulations by an Earth System Model to constrain parameters of an Earth System Model of Intermediate Complexity (EMIC).

2. Methods and Model Simulations

The EMIC used to obtain distributions of climate system properties is the Bern2.5D model. It is based on a zonally averaged dynamic ocean model and resolves the ocean basins of the Atlantic, Pacific, Indian, and Southern Oceans. The ocean model is coupled to a zonally and vertically averaged energy and moisture balance model of the atmosphere [*Schmittner and Stocker, 1999; Stocker et al., 1992*]. We employ the historical natural and anthropogenic annual mean radiative forcings of the Representative Concentration Pathways (RCP) scenarios to drive the model. We use the *20c3m* scenario during the period 1765–2005 and the RCP8.5 scenario onward for the 21st century.

The implementation and prior distributions of scaling factors in the Bern2.5D climate model, which account for the uncertainty in different forcing agents and climate system properties, are described in earlier studies [*Huber and Knutti, 2012; Knutti et al., 2003; Tomassini et al., 2007*]: in total, 13 parameters are sampled in the model: 3 physical parameters including the equilibrium climate sensitivity as well as 10 forcing scaling parameters accounting for forcing uncertainty, e.g., of greenhouse gases, direct and indirect aerosol effects, volcanic eruptions, and solar variations. The default parameter sampling setup uses a uniform distribution of ECS between 1 and 10°C. To test the sensitivity of our results to the choice of ECS prior distribution, we also computed all the results for the two nonuniform gamma distributions: a rather low-sensitivity distribution with shape and scale parameters of 2.1 and 0.95, resulting in a mean of 2°C (0.4–4.7°C, 5–95% range), as well as a larger sensitivity distribution with parameters of 4.3 and 0.95 with a corresponding mean of 4.1°C (1.5–7.8°C, 5–95% range) (see supporting information). The 13 model parameters are constrained with a Markov Chain Monte Carlo algorithm package for MATLAB [*Haario et al., 2006*]. Unforced internal variability is represented with an autoregressive process [*Tomassini et al., 2007*]. In order to increase the computational efficiency of the parameter-constraining process, the full Bern2.5D model is replaced with a neural network. We use 5000 independent time series in which the 13 model parameters are sampled to train a three-layer feedforward neural network built of 10 nodes [*Huber and Knutti, 2012*].

The 20-member initial condition ensemble is performed with the Community Earth System Model (CESM) version 1.0.4. The model consists of the Community Atmosphere Model version 4 (CAM4) and a fully coupled ocean, sea ice, and land surface components [*Gent et al., 2011*] with an ECS value of 4.1°C [*Meehl et al., 2013b*]. The transient climate response (TCR) is estimated to be 1.7°C based on a 10-member ensemble of a transient 1% increase in CO₂ scenario. The model is driven by historical forcing until 2005 and the RCP8.5 scenario until 2100. In order to create an initial condition member ensemble, a small random perturbation in the order of 10⁻¹³ is imposed on 1 January 1950, to the atmospheric initial condition field of the reference run, producing a 20-member ensemble covering the period 1950–2100.

As observations, we use the HadCRUT4 [*Morice et al., 2012*] data set for surface air temperature (*Had* in this paper) and the ocean heat uptake to 700 m data of *Domingues et al.* [2008] (DOM) and *Levitus et al.* [2012] (LEV). The reduced complexity model is constrained to both global temperature and ocean heat uptake to 700 m observations, resulting in the two combinations: *HadDOM* and *HadLEV*, respectively. When the reduced complexity model is constrained to the 20-member ensemble, the observations are replaced by the output of the simulations of the CESM climate model.

3. Results

The range of distributions for the equilibrium climate sensitivity and the transient climate response when all simulated observations by the Earth System Model are used until the year 2012 is shown in Figure 1. Allowing for larger uncertainties in ocean heat uptake observations (*HadDOM*, Figure 1a) leads to both higher ECS and TCR values compared with the setup where ocean uncertainties are small and almost constant (*HadLEV*, Figure 1b). The effect of internal variability is about as twice as large for ECS than for TCR: the mean values for ECS vary about 0.5°C (*HadLEV*) to 0.7°C (*HadDOM*) (1 σ), with similar values of 0.4°C and 0.8°C for the most likely values. This spread in the estimates corresponds to about 15% ($\pm 1\sigma$)

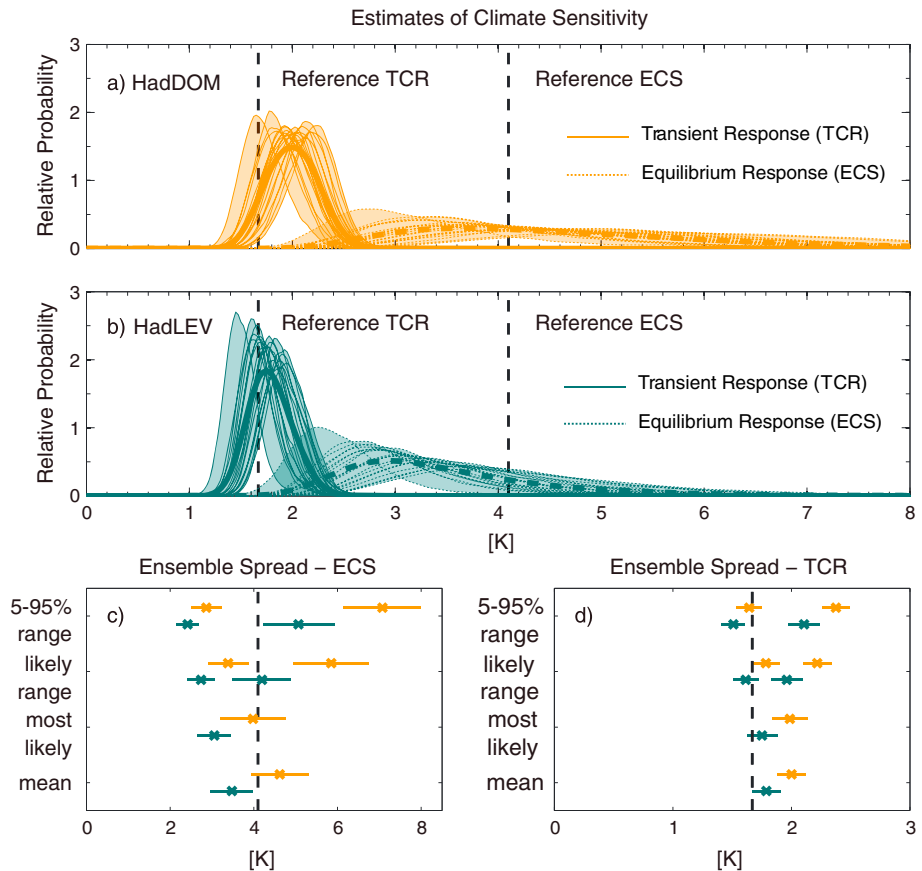


Figure 1. Estimates of climate sensitivity and the transient climate response: (a and b) Posterior distributions for equilibrium climate sensitivity (ECS) and transient climate response (TCR) are derived by constraining a reduced complexity model to the simulations of a 20-member initial condition ensemble of the Community Earth System Model with known sensitivities of 4.1°C (ECS) and 1.7°C (TCR) (black dashed line), which we treat as observations, using a uniform prior (see Methods). Values are shown for the two setups: Figure 1a shows HadDOM and Figure 1b shows HadLEV, which differ in their magnitude and treatment of uncertainty in ocean heat uptake observations (see Methods). Each curve represents an estimate for one realization of natural variability. (c and d) The ensemble spread and percentile ranges for ECS and TCR.

(13–20% for the most likely values) around the ensemble average. The TCR values vary in the order of about 0.1°C, corresponding to a spread of about 6% ($\pm 1\sigma$).

In contrast to the real world, this setup offers the opportunity to evaluate the probabilistic estimates obtained with this method since ECS and TCR of the Earth System Model are known (dashed lines in Figure 1). Within the setup of this study, the situation where only one realization of the climate system is available and which we are faced with in the real world has roughly about a 15% chance (HadDOM), respectively, a 45% chance (HadLEV), that the underlying sensitivity could lie outside the estimated likely range, compared to the expected 34% in a likely range (central 66%). Figure 1 shows that the HadDOM setup generally overestimates TCR, whereas ECS is underestimated with the HadLEV setup.

The joint evolution of ECS and TCR over time since the year 1970 is shown in Figure 2 both for estimates obtained from real-world observations and the corresponding simulations of the Earth System Model. We emphasize that the ECS of the real world and of the Earth System Model are different; thus, the figure emphasizes the temporal behavior of the two setups. The high and constant sensitivities between 1970 and 2000 are strongly related to the choice of prior distribution for ECS. For all three prior choices accounted for in this study, the ECS and TCR distributions show almost no change between 1970 and 2000, after which they start to converge to very similar values today (see supporting information).

The most striking difference between the observed and perfect-model setup can be seen since about the year 2000, where climate sensitivity estimates obtained from the observations decrease much sharper than

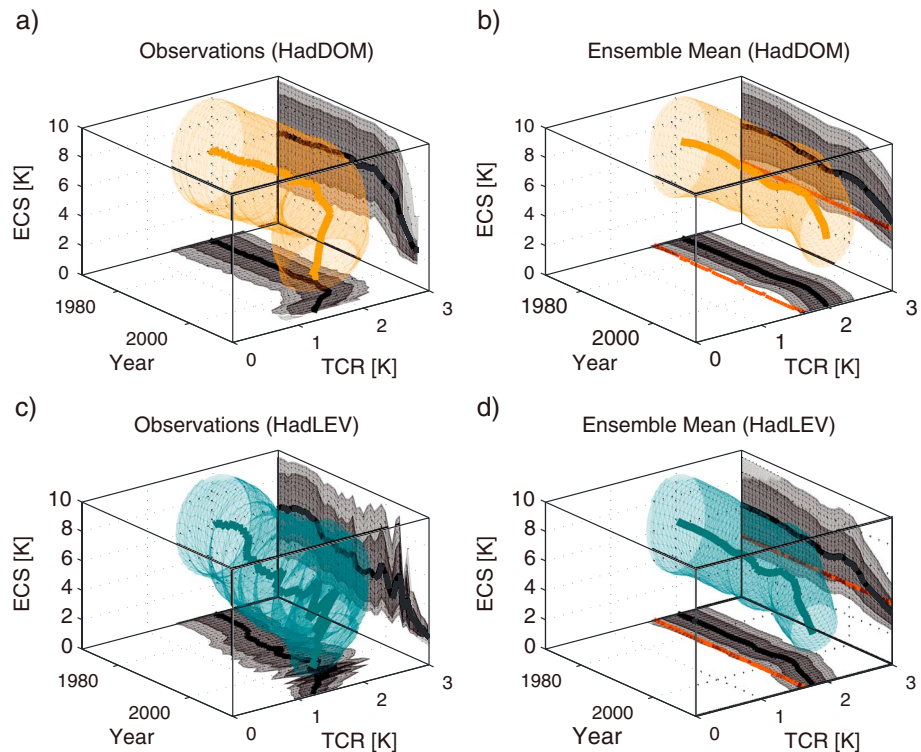


Figure 2. Temporal evolution of climate sensitivity and transient response estimates as a function of the calendar year: The colored tubes denote the joint 5–95% uncertainty range of the estimates of equilibrium climate sensitivity (ECS) and transient climate response (TCR), whereas the side projections show the likely (central 66%) and 5–95% range for two quantities individually. The solid lines denote the mean estimate. (a and c) The distributions where the reduced complexity climate model is constrained to two observational data sets (HadDOM and HadLEV), whereas (b and d) the simulations of the Earth System Model are taken as observational constraints. Figures 2b and 2d show the ensemble-averaged sensitivity distributions. The reference values of the ESM are denoted as red lines.

the corresponding values of the perfect-model case. Additionally, climate sensitivity can strongly vary between subsequent years when only small errors in ocean heat uptake are considered (HadLEV), suggesting that the reduced complexity model tries to fit interannual or decadal variations that cannot be reproduced by the model without variability.

The constrained climate model parameters allow us to compute probabilistic estimates of future temperature change under the RCP8.5 scenario that are consistent with past observed changes in global temperature and ocean heat uptake. Figures 3a and 3b show the ensemble-mean projections over time which generally agree well with the projected temperature change of the Earth System Model. Since the HadDOM setup slightly overestimates the reference TCR of the Earth System Model, it overestimates the warming toward the end of the 21st century.

The estimates presented here are Bayesian in nature, and their relation to commonly used ensemble-frequentist based skill score metrics such as in weather prediction has been neither clear nor possible owing to the limited number of model simulations and the lack of real-world observations for validation. The 20-member ensemble of the fully coupled model allows to compare the two different prediction methods for the first time and to answer, for example, whether at least 10 out of the 20 predictions fall within the central 50% prediction range derived with the reduced complexity climate model.

This concept is illustrated in Figures 3c and 3d for the decadal temperature projections over the 21st century. For each of the 90 decades starting from the 2000s, we have 20 “true” values as simulated by the Earth System Model (ESM) as well as 20 corresponding probabilistic distributions for the two setups derived with the EMIC. Starting with the first ensemble member, we can compute in which predicted central range (i.e., 5–95%, likely range) the true ESM value lies. Thus, in the case where prediction ranges can be interpreted in an ensemble-frequentist sense, 66% of the cases should contain the true simulated value within the estimated likely range

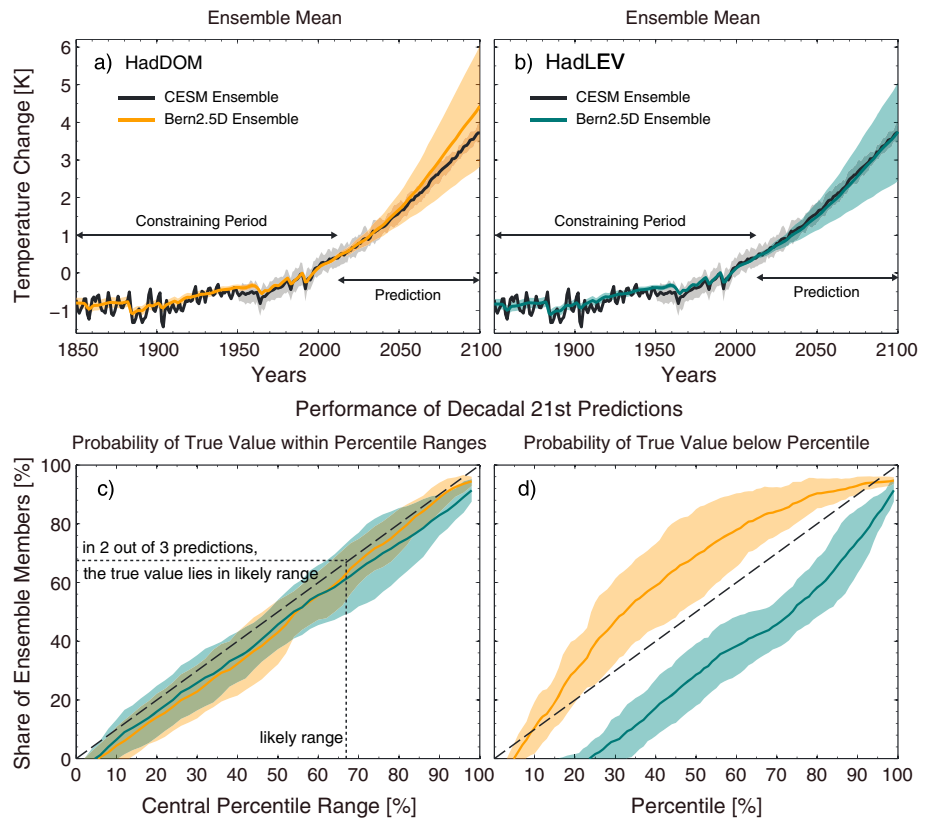


Figure 3. Evaluation of probabilistic temperature projections for the 21st century: Projected temperature change under the RCP8.5 emission scenario relative to the reference period 1986–2005 for the HadDOM and HadLEV setup. (a and b) The ensemble-averaged projections for the 20-member ensemble of the Earth System Model (black solid lines) and for the reduced complexity climate model (colored solid lines) that is constrained to the simulations of global temperature and upper ocean heat uptake of the Earth System Model. Shaded regions denote the $\pm 2\sigma$ uncertainty range. (c and d) The skill of the decadal predictions of future temperature change during the 21st century in terms of the number of cases in which the true simulated value by the Earth System Model falls within a specific central percentile range (Figure 3c) or below a specific percentile (Figure 3d). The black dashed line corresponds to the linear relationship where, e.g., in two out of three (66%) of the predictions by the EMIC include the true value of the ESM within the predicted likely range (central 66%). Figures 3c and 3d show the average (solid line) and standard deviation (shading) across all 21st century decadal predictions for a given percentile.

(defined as central 66% of the distribution predicted by the EMIC), defining a linear relation against which our predictions can be tested (dashed black lines in Figures 3c and 3d). For each of the decades, we obtain 90 estimates of the percentage of the cases containing the true simulated value, and Figure 3c shows that on average, such a linear relation is indeed found.

However, the result is not as robust for a specific lower (or upper) bound compared to the central prediction range as depicted in Figure 3d. For example, the figure shows that the number of cases, in which the true reference values lie below the 20% percentile of the prediction estimate is overestimated in the HadLEV case, whereas it is underestimated in the HadDOM case due to the skewed sensitivity distributions and their biases relative to the reference sensitivity shown in Figure 1.

4. Summary and Conclusions

Using a probabilistic setup of a reduced complexity model and an ensemble of an Earth System Model, we showed that unforced climate variability is important in the estimation of the climate sensitivity, in particular when estimating the most likely value, and more so for the equilibrium than for the transient response. A particular emphasis was put on the role of uncertainties of upper ocean heat uptake observations by taking two different observational data sets into account. The spread of climate sensitivity estimates presented in Figure 1 suggests that the effect of the treatment of these uncertainties on climate sensitivity estimate is of

similar order than that of internal variability: the difference in the ensemble mean estimate of ECS of the two setups of about 1.1°C is even larger than the spread for the HadDOM ensemble of 0.7°C.

The evaluation of the probabilistic temperature projections derived with the reduced complexity model with the 21st century simulations of the Earth System Model showed a strong agreement in terms of the ensemble mean values. In this setup, projections of individual ensemble members are more robust when central ranges are considered than individual prediction percentiles (i.e., the median), since they are either high (HadDOM) or low (HadLEV) biased depending on the two setups. In addition, the effect of the spread and bias in the transient response on future temperature projections increases with lead time since the estimate of TCR with this method is strongly related to the total radiative forcing. For example, a change in TCR of 0.5°C corresponds to a change in the mid-21st century decadal temperature prediction of ~0.5°C, whereas it amounts to three times as much toward the end of the century.

This study also highlights the importance of the last decade in estimating climate sensitivity, in particular of the role of ocean heat uptake: the climate signal gets stronger and emerges from noise, and uncertainties are reduced, which helps to constrain model parameters. But at the same time variability that is not reproduced in the model becomes important, and structural biases of simple models become evident.

Overall, we find a delicate balance between a poor constraint and, as a consequence, issues with the choice of priors as long as the data is not very clear (as seen in many early studies on this topic), and overfitting and structural problems of simple models once the data constraint becomes strong. The framework employed here provides a powerful test of such methods, but at the same time raises questions about the reliability of the results, in particular when oversimplified models are used, and when internal variability is not carefully accounted for.

Acknowledgments

The Editor thanks an anonymous reviewer for assisting in the evaluation of this paper.

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(3), 253–271.
- 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 multi-decadal sea-level rise, *Nature*, 453(7198), 1090–1096.
- 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(5552), 113–117.
- 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. Atmos.*, 118(3), 1139–1150.
- Gent, P. R., et al. (2011), The Community Climate System Model Version 4, *J. Clim.*, 24(19), 4973–4991.
- Guemas, V., F. J. Doblas-Reyes, I. Andreu-Burillo, and M. Asif (2013), Retrospective prediction of the global warming slowdown in the past decade, *Nature Clim. Change*, 3(7), 649–653.
- Haario, H., M. Laine, A. Mira, and E. Saksman (2006), DRAM: Efficient adaptive MCMC, *Stat. Comput.*, 16, (4), 339–354.
- Huber, M., and R. Knutti (2012), Anthropogenic and natural warming inferred from changes in Earth's energy balance, *Nat. Geosci.*, 5(1), 31–36.
- IPCC (2013), Summary for Policymakers, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. F. Stocker et al., Cambridge Univ. Press, Cambridge, U. K., and New York.
- Knutti, R., and G. C. Hegerl (2008), The equilibrium sensitivity of the Earth's temperature to radiation changes, *Nat. Geosci.*, 1(11), 735–743.
- 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(6882), 719–723.
- Knutti, R., T. F. Stocker, F. Joos, and G. K. Plattner (2003), Probabilistic climate change projections using neural networks, *Clim. Dyn.*, 21(3–4), 257–272.
- Kosaka, Y., and S. P. Xie (2013), Recent global-warming hiatus tied to equatorial Pacific surface cooling, *Nature*, 501(7467), 403–407.
- Levitus, S., et al. (2012), World ocean heat content and thermocline sea level change (0–2000 m), 1955–2010, *Geophys. Res. Lett.*, 39, L10603, doi:10.1029/2012GL051106.
- Lewis, N. (2013), An objective Bayesian, improved approach for applying optimal fingerprint techniques to estimate climate sensitivity, *J. Clim.*, 26(19), 7414–7249.
- 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, *Nat. Geosci.*, 5(2), 110–113.
- Meehl, G. A., J. M. Arblaster, J. T. Fasullo, A. X. Hu, and K. E. Trenberth (2011), Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods, *Nat. Clim. Change*, 1(7), 360–364.
- Meehl, G. A., A. X. Hu, J. M. Arblaster, J. Fasullo, and K. E. Trenberth (2013a), Externally Forced and Internally Generated Decadal Climate Variability Associated with the Interdecadal Pacific Oscillation, *J. Clim.*, 26(18), 7298–7310.
- Meehl, G. A., W. M. Washington, J. M. Arblaster, A. X. Hu, H. Y. Teng, J. E. Kay, A. Gettelman, D. M. Lawrence, B. M. Sanderson, and W. G. Strand (2013b), Climate Change Projections in CESM1(CAM5) Compared to CCSM4, *J. Clim.*, 26(17), 6287–6308.
- Morice, C. P., J. J. Kennedy, N. A. Rayner, and P. D. Jones (2012), Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set, *J. Geophys. Res.*, 117, D08101, doi:10.1029/2011JD017187.
- Olson, R., R. Sriver, W. Chang, M. Haran, N. M. Urban, and K. Keller (2013), What is the effect of unresolved internal climate variability on climate sensitivity estimates?, *J. Geophys. Res. Atmos.*, 118, 4348–4358, doi:10.1002/jgrd.50390.
- Otto, A., et al. (2013), Energy budget constraints on climate response, *Nat. Geosci.*, 6(6), 415–416.
- Rohling, E. J., et al. (2012), Making sense of palaeoclimate sensitivity, *Nature*, 491(7426), 683–691.
- Schmittner, A., and T. F. Stocker (1999), The stability of the thermohaline circulation in global warming experiments, *J. Clim.*, 12(4), 1117–1133.

- Stephens, G. L., J. L. Li, M. Wild, C. A. Clayson, N. Loeb, S. Kato, T. L'Ecuyer, P. W. Stackhouse, M. Lebsock, and T. Andrews (2012), An update on Earth's energy balance in light of the latest global observations, *Nat. Geosci.*, *5*(10), 691–696.
- Stocker, T. F., D. G. Wright, and L. A. Mysak (1992), A zonally averaged, coupled ocean atmosphere model for paleoclimate studies, *J. Clim.*, *5*(8), 773–797.
- 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*(7), 1239–1254.