

Constraining Climate Sensitivity from the Seasonal Cycle in Surface Temperature

RETO KNUTTI AND GERALD A. MEEHL

National Center for Atmospheric Research, Boulder, Colorado*

MYLES R. ALLEN AND DAVID A. STAINFORTH

Atmospheric and Oceanic Physics, Oxford University, Oxford, United Kingdom

(Manuscript received 16 June 2005, in final form 29 November 2005)

ABSTRACT

The estimated range of climate sensitivity has remained unchanged for decades, resulting in large uncertainties in long-term projections of future climate under increased greenhouse gas concentrations. Here the multi-thousand-member ensemble of climate model simulations from the climateprediction.net project and a neural network are used to establish a relation between climate sensitivity and the amplitude of the seasonal cycle in regional temperature. Most models with high sensitivities are found to overestimate the seasonal cycle compared to observations. A probability density function for climate sensitivity is then calculated from the present-day seasonal cycle in reanalysis and instrumental datasets. Subject to a number of assumptions on the models and datasets used, it is found that climate sensitivity is very unlikely (5% probability) to be either below 1.5–2 K or above about 5–6.5 K, with the best agreement found for sensitivities between 3 and 3.5 K. This range is narrower than most probabilistic estimates derived from the observed twentieth-century warming. The current generation of general circulation models are within that range but do not sample the highest values.

1. Introduction

Projections of future increases in global temperature for a given emission scenario are uncertain, with recent quantitative methods estimating a spread of the models of about 30% around the best-guess projection (Knutti et al. 2002; Stott and Kettleborough 2002). This spread arises mainly from uncertainties in the carbon cycle, radiative forcing for given atmospheric concentrations, climate sensitivity, and how the ocean mixing affects the transient ocean heat uptake. The current range of climate sensitivity is the dominant contribution to the total uncertainty in future warming projections after about the year 2050 (Knutti et al. 2002). Climate sensitivity is usually referred to as the global mean equilibrium near-surface warming for doubling the atmo-

spheric CO₂ concentration, equivalent to a radiative forcing of about 3.7 W m⁻² (Myhre et al. 1998). It is used to quantify all important feedbacks, including albedo, cloud, and water vapor feedbacks, which potentially amplify the warming. Both global temperature increase and its uncertainty on time scales of centuries scale almost linearly with sensitivity.

The almost canonical range of 1.5–4.5 K for climate sensitivity was derived from the range covered by different atmosphere–ocean general circulation models (AOGCMs) and has remained virtually unchanged for decades (Charney 1979; Houghton et al. 1996, 2001). No probability was ever formally attached to this range, but sometimes assumed ad hoc due to the lack of better knowledge (Wigley and Raper 2001). However, if we are interested for example in the likelihood of exceeding a certain temperature threshold for a given forcing, a probability density function (PDF) for sensitivity is inevitably needed.

A number of methods were proposed to constrain climate sensitivity from observations in a probabilistic way. Studies based on a combination of the reconstructed radiative forcing and the observed surface warming and ocean heat uptake (Andronova and

* The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Corresponding author address: Reto Knutti, NCAR, P.O. Box 3000, Boulder, CO 80307.
E-mail: knutti@ucar.edu

Schlesinger 2001; Forest et al. 2002; Gregory et al. 2002; Knutti et al. 2002, 2003; Forest et al. 2006; Frame et al. 2005) find ranges of sensitivity consistent with observations that are much larger than the range of 1.5–4.5 K given by the Intergovernmental Panel on Climate Change (IPCC) Third Assessment Report (TAR; Houghton et al. 2001). In particular, the upper end is difficult to constrain, and several studies have highlighted the possibility of sensitivity being 8 K or more. Ensembles of climate models with perturbed physics evaluated against present-day climatology also yield wide sensitivity ranges (Murphy et al. 2004; Stainforth et al. 2005). Paleoclimate data seems to provide some constraints on climate sensitivity (Hoffert and Covey 1992; Lea 2004) but attempts are hampered by an often poor quantification of the uncertainties associated with paleoclimate data, and more recent estimates based on climate models (Annan et al. 2005; Hegerl et al. 2006; Schneider von Deimling et al. 2006) are not conclusive. The fact that only proxies of temperature are available and that those can be found only in some locations, makes expert judgement inevitable. Thus, it seems that recent quantitative attempts to constrain climate sensitivity have rather increased than decreased the likely range compared to IPCC TAR (Houghton et al. 2001). In particular, the low probability but high impact case of very large sensitivities is a reason for concern in adaptation or mitigation planning of future climate change.

In this paper we propose that the amplitude of the seasonal cycle in temperature and its geographical distribution provides a constraint on climate sensitivity and can thus be used for the evaluation of model performance. The idea is based on the fact that seasonal sensitivity (i.e., the local temperature change for the solar insolation difference between summer and winter at a given location) is at least partly determined by the same processes as climate sensitivity, the equilibrium global temperature response to a global change in forcing. For some feedbacks, it is obvious that their strength on a seasonal scale will be relevant for global warming (e.g., how fast the land warms or cools compared to the ocean). Hall and Qu (2006) also found a strong relationship of the surface albedo feedback on seasonal and on decadal time scales for a series of AOGCMs. In this case, a strong dependence of the surface albedo and snow cover on temperature results in both a strong seasonal cycle and a strong global warming signal, in particular in the high-latitude regions. This is in line with the fact that we find high-latitude Northern Hemisphere regions to provide the strongest constraint on climate sensitivity. For other feedbacks like water vapor and clouds, it is not clear how seasonal and decadal

effects are related. Patterns of temperature differences like the land–ocean contrast, Northern Hemisphere meridional temperature contrast, interhemispheric temperature contrast, the seasonal cycle, and the diurnal temperature range have been found to correlate strongly with global temperature (Braganza et al. 2003), suggesting also that the amplitude of those patterns is determined at least partly by the same feedbacks as those that control global temperature. A full quantification of all individual relevant feedbacks is impossible at this stage, but also not needed. We simply argue that there is (at least in some regions) a relation between the amplitude of the seasonal temperature cycle and climate sensitivity if climate model parameters are varied, and that such a relation can be approximated by a statistical method. Using that relation and observations on the amplitude of the seasonal cycle allows us to derive a PDF of climate sensitivity.

2. Model and statistical procedure

Climateprediction.net (CPDN) is a distributed computing project that runs thousands of climate model simulations on standard public or home computers (Stainforth et al. 2005; more information available online at <http://climateprediction.net/>). Parameters and initial conditions in a version of the Third Hadley Centre atmosphere-slab ocean model (HadSM3; Pope et al. 2000; Murphy et al. 2004) are perturbed to explore the widest possible range of model responses to doubling atmospheric CO₂. First results in this simplified setup yielded a wide range of responses, with climate sensitivity values from less than 2 to more than 11 K (Stainforth et al. 2005). Even model versions with sensitivities larger than 10 K could not be excluded based on annual mean climatology.

For the analysis presented here, 2500 CPDN simulations were used. Quality control and the calculation of climate sensitivity were done as described by Stainforth et al. (2005). Seasonal amplitude (S) in temperature was calculated as boreal summer June–August (JJA) minus winter December–February (DJF) temperature in 26 regions, the same as used in previous studies (Giorgi and Francisco 2000; Houghton et al. 2001) plus hemispheric and extratropical hemispheric averages (see Fig. 1). Mean values were calculated from years 7–15 for both the control and $2 \times \text{CO}_2$ phase and were taken over the full latitude–longitude box for each region, and thus can include some ocean area (see Stainforth et al. 2005 for details about the experimental setup and available output from the CPDN project).

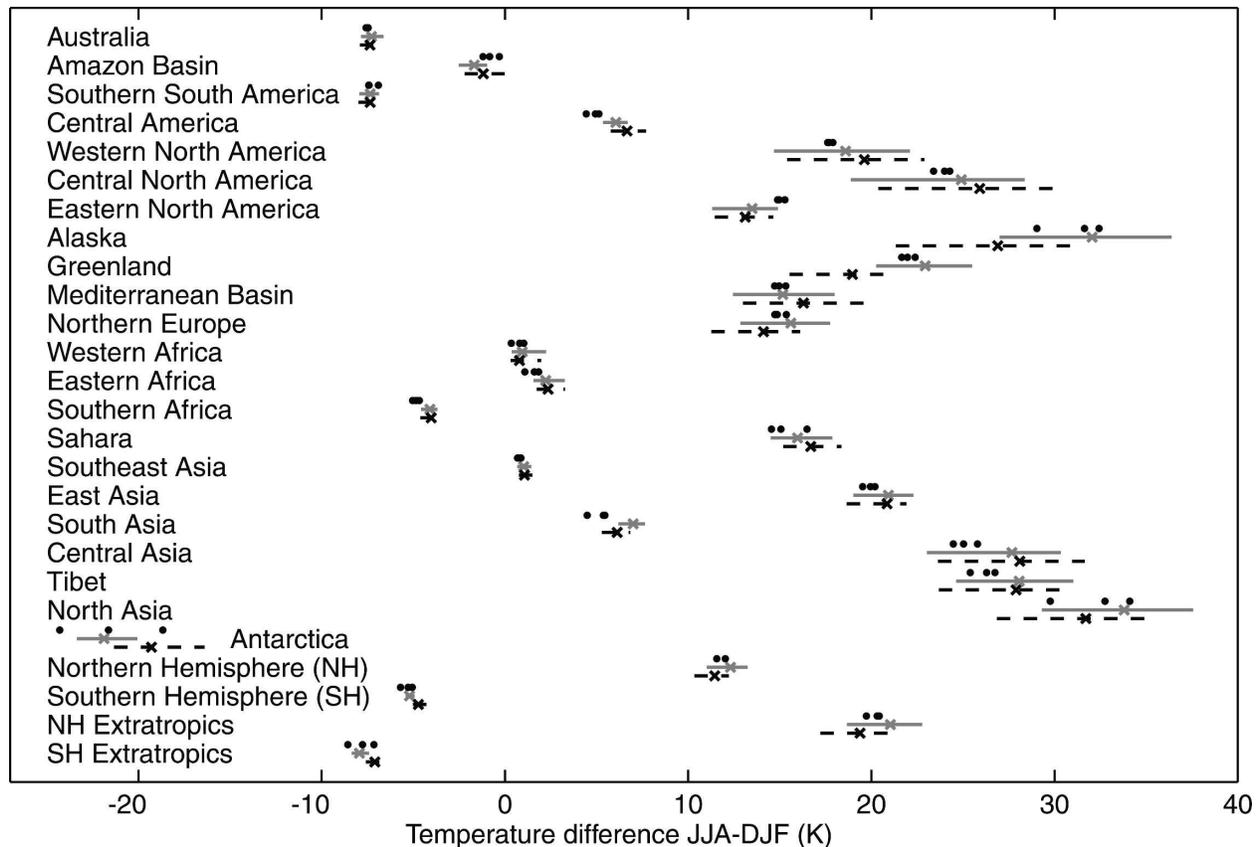


FIG. 1. The 5%–95% ranges (lines) and medians (crosses) of the boreal summer – winter (JJA – DJF) temperature covered by all CPDN simulations, for the control (gray solid) and $2 \times \text{CO}_2$ simulations (black dashed) and for each region considered. Observed mean values for the ERA-40, NCEP–NCAR, and HadCruT2v datasets are given as black dots for comparison.

The 5%–95% range for S_{co} in the control phase across the ensemble is shown in Fig. 1 (gray solid lines) for each region. It is important to note that seasonality is expected to change in a warming world (e.g., in high northern latitudes, winter temperatures warm faster than summer temperatures). The 5%–95% range for S_{2x} in the $2 \times \text{CO}_2$ run is given as black dashed lines in Fig. 1 for comparison. The observed seasonality is thus neither representative of a preindustrial control nor a $2 \times \text{CO}_2$ simulation. We therefore assume here that seasonality changes linearly with global temperature in each model version, and to be consistent with observations, interpolate seasonality for present-day climate S_{pr} linearly between the control and $2 \times \text{CO}_2$ case for each model version using the effectively realized global temperature increase of years 7–15 in the $2 \times \text{CO}_2$ case relative to the control and assuming 0.35-K warming of the period 1950–2000 compared to the preindustrial period. The effect of using this present-day amplitude of the seasonal cycle instead of S_{co} from the control phase is small and does not change any of the conclusions. An example of the interpolated present-day amplitude of

the seasonal cycle in temperature S_{pr} over four regions as a function of climate sensitivity from 2500 simulations of the CPDN dataset is shown in Fig. 2. Averages over the years 1950–2000 obtained from all forcing simulations of the twentieth-century simulations from 17 coupled AOGCMs (available online at http://www-pcmdi.llnl.gov/ipcc/about_ipcc.php) participating in the IPCC Fourth Assessment Report (AR4) are shown as circles for comparison. The mean seasonal cycle from the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40; Simmons and Gibson 2000) and the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR; Kalnay et al. 1996) reanalysis datasets and from the Hadley Centre/Climatic Research Unit instrumental dataset (HadCruT2v; Jones and Moberg 2003) are shown as horizontal lines for comparison. In most of the extratropical Northern Hemisphere regions (e.g., Figs. 2a,b), there is significant positive correlation, with the low sensitivity model versions underestimating and the very high sensitivity model versions strongly overestimating the observed

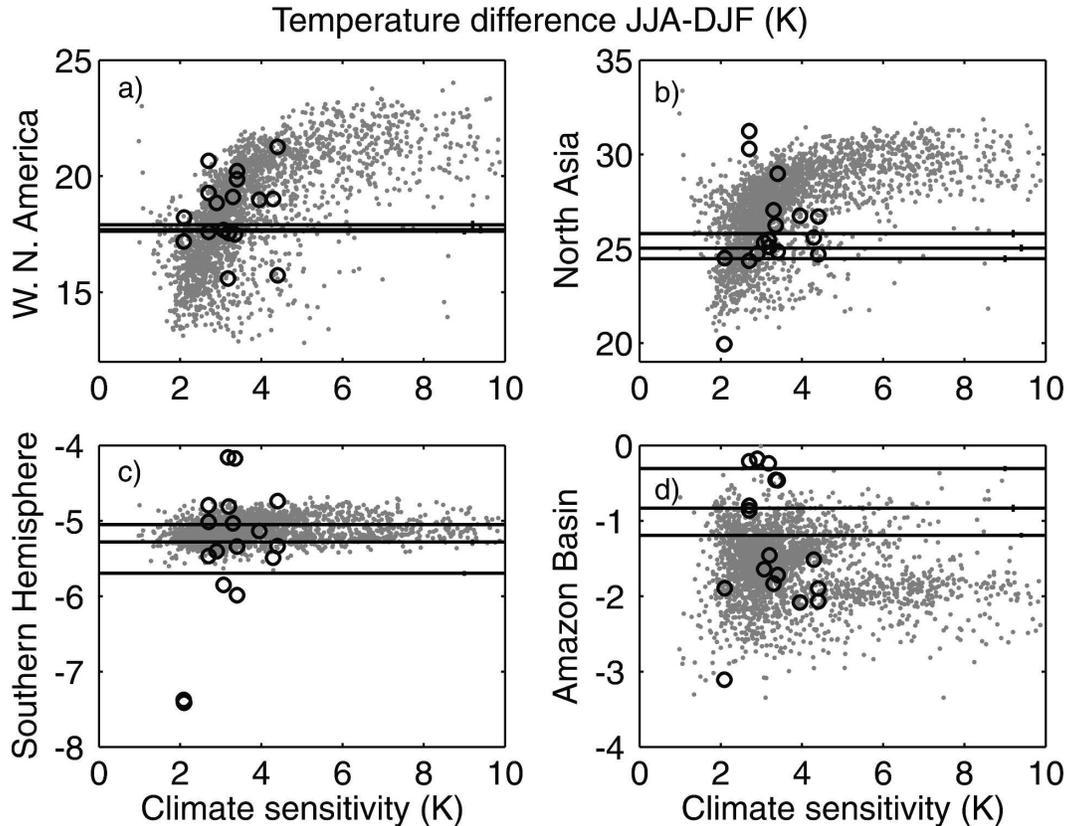


FIG. 2. The amplitude of the seasonal cycle in near-surface temperature in (a) western North America, (b) north Asia, (c) the SH, and (d) the Amazon basin vs climate sensitivity. Each gray dot represents a simulation from the CPDN project. Black horizontal lines mark the means from the ERA-40, NCEP-NCAR, and HadCruT2v datasets over the same region. The uncertainties in each of the mean climatologies are smaller than the spread of the three datasets. Black circles mark amplitudes of the seasonal cycle in the IPCC twentieth-century simulations.

seasonality. Correlation varies strongly by region, with some regions showing no correlation at all or showing a rather complicated pattern (e.g., Figs. 2c,d). From Fig. 2c and Fig. 1 it becomes clear that seasonality does not change at all or is not correlated with sensitivity in some regions, and thus does not provide any constraint.

To find the relation between climate sensitivity and the seasonality S_{pr} in different regions, we use a neural network. Neural networks are used to solve a variety of problems in pattern recognition and classification, tracking, control systems, fault detection, data compression, feature extraction, signal processing, optimization problems, associative memory, and more. There are many different types of neural networks, each suitable for specific applications.

A feed-forward neural network with 10 neurons and sigmoid transfer function in the input layer and one neuron with linear transfer function in the output layer is used here to predict climate sensitivity from the seasonality S_{pr} . The Levenberg–Marquardt algorithm (Hagan and Menhaj 1994) is the most efficient learning

algorithm for this application. The choice of the network type and size as well as the learning algorithm depend on the problem considered. A detailed introduction to neural network architectures, learning rules, training methods, and applications was written by Hagan et al. (1996). In this study, 60% of the simulations are used for training, the rest is used for validation to test against overfitting. First, the neural network is trained using a random subset of simulations as the training set. Essentially, this procedure minimizes the error between predicted and true sensitivity in the training set by adjusting the weights in each neuron. Thus, instead of following a specified set of rules, the neural network learns the underlying input–output relationship between the seasonality (input) and climate sensitivity (output) from the examples presented in the training set. If training is successful, the neural network will later be able to predict the climate sensitivity values for a pattern of seasonality to the extent the latter contains information about the former. The training ensures that the regions that have a strong relation to

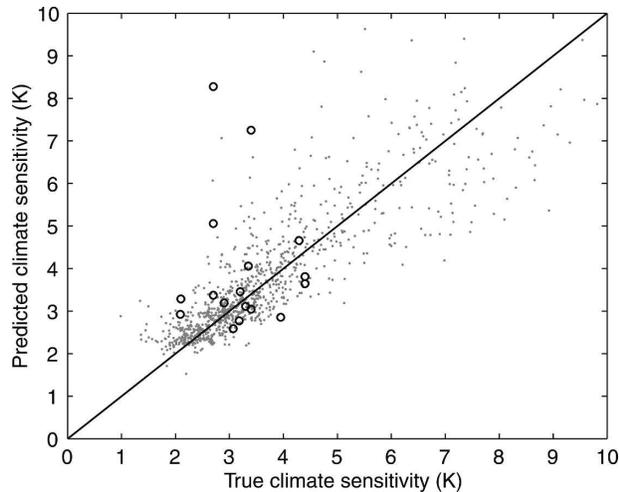


FIG. 3. Climate sensitivity predicted with the neural network from seasonality of temperature vs true sensitivity from a subset of 1000 CPDN simulations not used for training (gray dots) and for the 17 IPCC AOGCMs (black circles). Correlation for the CPDN simulations is $r = 0.88$.

sensitivity are given more weight than those that have a weak relation. Decision neurons (sigmoid transfer functions) effectively lead to a piecewise approximation of the relation in parameter space. In the absence of noise, if enough information in the training set and sufficient neurons and layers are provided, a neural network is able to fit any relation, nonlinear or even discontinuous, to arbitrary accuracy. The remaining 40% of the simulations are used to prevent overfitting by an early stopping mechanism, but overfitting is not a problem here. Overfitting refers to the situation when the neural network has too many degrees of freedom and starts to approximate the noise in the training subset, which is not robust throughout the whole dataset. Checking the neural network in each training step against an independent set of simulations allows the detection of this problem and training can be stopped if overfitting starts to occur. For a given training set, multiple training attempts are done with random initial weights in the neurons to ensure the best possible performance, measured as rms difference between predicted and true climate sensitivity on the rest of the simulations not used for training. The predicted versus true sensitivities of an independent set of simulations not used for training are shown in Fig. 3. In general, we find good agreement, with a correlation coefficient $r = 0.88$. However, the agreement is far from perfect, indicating that the seasonality contains not all information needed and is only partly successful in predicting climate sensitivity. The mean absolute error (predicted minus true sensitivity) is about 0.5 K for sensitivities around 3 K, and increases

to about 1.5 K for sensitivities larger than 8 K. The distribution of the error ε between predicted and true sensitivity is assumed to be normal (but dependant on sensitivity) and will be used in the next steps. Second, NCEP–NCAR and ERA-40 reanalyses and the HadCruT2v instrumental datasets are used to calculate a climatology and its uncertainty due to interannual variability (the uncertainty of the mean climatology is the standard deviation of the individual years divided by the square root of the number of years). Ten thousand sets of the 26 regional amplitudes of the seasonal temperature cycle are then generated assuming a normal distribution of the uncertainty in the climatology of each region. Third, each of these perturbed climatology sets is taken as an input into the neural network, predicting one value for climate sensitivity. A random error is added to the predicted sensitivity according to the error ε estimated from simulations with similar sensitivity in the CPDN validation set. Fourth, this procedure is repeated 50 times using other random training subsets, accounting for the fact that the neural network training will not always converge to the same solution when using a different training set.

The final probability density functions is therefore a composite of 500 000 simulated climate sensitivity values. It takes into account the uncertainty due to interannual variability in the climatological dataset and the fact that the neural network is not a perfect predictor, and it uses both different subsets for training and multiple initial values for the weights in the neurons. The total uncertainty in the predicted climate sensitivity depends on the uncertainty assumed for the mean climatology. The assumptions made for the observational uncertainty as well as the assumptions in this statistical method are discussed in a separate section below.

3. Results and discussion

The median and 5%–95% ranges for climate sensitivity derived from observations and reanalysis climatology are shown in Fig. 4. Median values for sensitivity using the land regions only are found at 3.2, 3.3, and 3.4 K for ERA-40, NCEP–NCAR, and HadCruT2v, respectively. Under the “zero structural uncertainty” assumption in which the only sources of error considered are the uncertainty in the neural network model and uncertainty in the true amplitude of the seasonal cycle due to internal climate variability, the lower (5%) bounds are 2.2, 2.2, and 2.3 K for ERA-40, NCEP–NCAR, and HadCruT2v, respectively. The upper (95%) bounds are 4.3, 4.6, and 4.4 K, respectively. The average of the three ranges is 2.2–4.4 K, the median is 3.3 K. The solid lines in Fig. 4 denote the PDFs in the

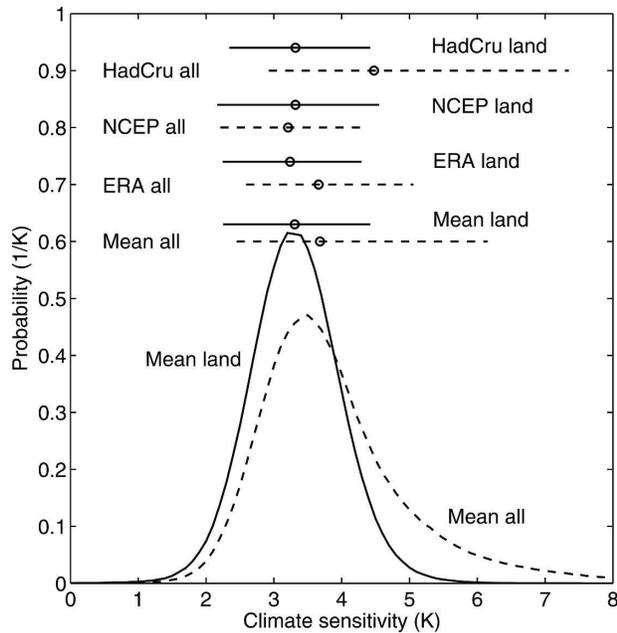


FIG. 4. PDFs for climate sensitivity derived from the ERA-40, NCEP–NCAR, and HadCruT2v datasets using all possible regions (“all”) or only those over land (“land”). Lines mark 5%–95% ranges; circles mark medians. The constraints are stronger if only the land regions are used (solid lines) compared to when hemispheric averages are included (dashed lines). The solid black PDF is the mean of the three PDFs using land regions only and is considered to be the standard case when neglecting structural uncertainties. The median in this case is 3.3 K; the 5%–95% range is 2.2–4.4 K.

standard case where all land regions were used, except for those regions where the climatological value was found to be outside the range covered by the CPDN dataset. Since there is a high degree of correlation between the seasonal temperature amplitude in different regions, omitting a few regions does not affect the fitting procedure. The dashed lines show the cases where the total Northern Hemisphere (NH) and the extratropical NH were used as an additional input. While the median and lower limit are similar, the upper limit is harder to constrain. We believe that the use of a slab ocean model in CPDN means that temperature over ocean follows the sea surface temperature dataset used in the spinup phase very closely, almost independent of the model parameters used, so the relationship (if any) between parameters and seasonal cycle over the ocean becomes arbitrary. This is supported by the fact that the climatology of the whole Southern Hemisphere (SH) and the extratropical SH of all model versions only covers a very small range, and that seasonality shows no correlation to climate sensitivity in these ocean dominated regions. Thus, Fig. 4 shows that the best guess and lower limit are quite robust for different assump-

tions and different climatologies, while the upper limit is more sensitive. Thus our most optimistic scenario of (unrealistically) assuming no structural uncertainty suggests that climate sensitivities below about 2 K and above 4.5 K can be ruled out at the 5% percentile, with a best guess around 3–3.5 K, in very good agreement with most AOGCMs. The lower bound is robust against all assumptions. This suggests strong evidence for a substantial net positive net feedback in the climate system and provides a lower bound on the expected climate change of the future. The upper bound, however, depends on how structural uncertainty is accounted for. This is further discussed in the section below.

4. Caveats

The results presented here depend on a number of assumptions. First, the HadSM3 model used in the CPDN project could have a bias in seasonality (i.e., over- or underestimate seasonality systematically in certain regions regardless of the parameters chosen), causing the PDF of sensitivity to be shifted by an unknown amount. Second, the neural network could fit a relation between seasonality and sensitivity that only exists in the HadSM3 model. As a weaker statement, one could argue that although there is a general relation between the two, the neural network might fit certain HadSM3 specific patterns that are not found in the real world, thus causing the uncertainty to be underestimated. Those uncertainties and all aspects of the model that cannot be varied by changing parameters (e.g., grid type and resolution, numerical schemes, forms of parameterizations, and processes resolved or neglected) are summarized here as “structural uncertainties.” They are difficult to estimate quantitatively; however, the following arguments point to the validity of the approach.

We provide an evaluation of the method by predicting climate sensitivity for other independent AOGCMs, using the exact same statistical procedure. In contrast to the real world, the sensitivity predicted from the seasonal cycle of an AOGCM can be compared to the true sensitivity obtained from a $2 \times \text{CO}_2$ simulation. Instead of reanalysis or observational data, output from coupled climate models from the years 1950–2000 is taken from transient twentieth-century simulations calculated for the upcoming IPCC AR4. The simulations include all radiative forcing components implemented by the modeling groups. Ensemble means are used for models where multiple simulations are available. The predicted median sensitivities and 5%–95% uncertainty ranges for the 17 models or model versions available are shown in Fig. 5, the predicted

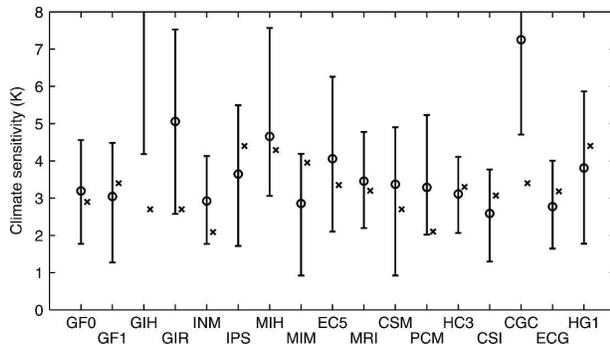


FIG. 5. Climate sensitivity predicted for 17 AOGCMs. Black circles indicate the median, black lines the 5%–95% range and black crosses mark the true model sensitivity. In the models, the amplitude of the seasonal cycle was calculated from 1950 to 2000 from the all-forcing simulations over the twentieth century to be comparable with the observational and reanalysis data. The models with their corresponding institutions are from left to right: the Geophysical Fluid Dynamics Laboratory Coupled Models Versions 2.0 (GFDL-CM2.0) and 2.1 (GFDL-CM2.1), the GISS-EH, GISS-ER, the Institute for Numerical Mathematics Coupled Model Version 3.0 (INM-CM3.0), the Institut Pierre Simon Laplace Coupled Model Version 4 (IPSL-CM4), the Model for Interdisciplinary Research on Climate Version 3.2 [MIROC3.2(hires)] and [MIROC3.2(medres)], the Max Planck Institute for Meteorology Ocean Model (ECHAM5/MPI-OM), the Meteorological Research Institute Coupled GCM Version 2.3.2 (MRI-CGCM2.3.2), the National Center for Atmospheric Research Community Coupled System Model Version 3 (CCSM3), the National Center for Atmospheric Research Parallel Climate Model (PCM), the Met Office (UKMO) HadCM3, the Commonwealth Scientific and Industrial Research Organisation Model (CSIRO-Mk3.0), the CCCma Coupled GCM Version 3.1 [CGCM3.1(T47)], the Meteorological Institute of the University of Bonn/Meteorological Research Institute of the Korea Meteorological Administration model (ECHO-G), and the UKMO Hadley Centre Global Environment Model (HadGEM1). (Model details and references can be found online at http://www-pcmdi.llnl.gov/ipcc/model_documentation/ipcc_model_documentation.php.)

medians versus true sensitivities are also shown as circles in the scatterplot in Fig. 3. In general, the predicted median sensitivity from the transient simulation is within less than 0.7 K (rms) of the true sensitivity calculated from the corresponding $2 \times \text{CO}_2$ slab model versions. Exceptions are the two National Aeronautics and Space Administration (NASA) Goddard Institute for Space Studies (GISS) and the Canadian CGCM models where the pattern of seasonality is unlike any of the CPDN simulations. In such a case, the neural network is unable to extrapolate from what it has been trained for. The three models are excluded for the rest of the discussion. Not surprising, the predicted uncertainty is smallest and agreement of the median with the true sensitivity is excellent for the coupled Third Hadley Centre Coupled Model (HadCM3), which shares the atmospheric component with the HadSM3 model

used by CPDN. We consider this to be a powerful test of our method. It shows that the relation between seasonality and sensitivity derived from slab control and $2 \times \text{CO}_2$ runs can be applied to output from a transient twentieth-century simulation. In a wider sense, this means that the results of a grand ensemble of a slab models can, to some extent, be used to predict some features of the results of transient, coupled simulations.

It could be argued that predicting sensitivity is inherently more difficult for another AOGCM than for the observations or reanalysis case. Although the reanalysis model also has deficiencies, it is forced to follow closely the observed data and is thus considered to be the “truth.” Predicting sensitivity for an AOGCM on the other hand has to deal with systematic errors in both the CPDN model HadSM3 and the independent model where sensitivity is to be predicted. On the other hand, many AOGCMs have errors and biases in common, so in the event of all climate-resolution GCMs sharing a common bias with respect to the real world, we appear to be doing misleadingly well in predicting GCM sensitivity. The fact that the error of the predicted sensitivity (0.7-K rms difference of predicted median minus true sensitivity for the 14 models) is smaller than the uncertainty range suggested by the neural network method (on average about 1.0 K, one standard deviation) indicates that the neural network method is not underestimating the uncertainty in predicting climate sensitivity due to structural uncertainties, as far as these are represented by intermodel differences. More encouraging still, there is no evidence of the neural network method over- or underestimating sensitivity in a systematic way. Thus, from the AOGCM test there is no evidence that we are underestimating the width of the sensitivity PDFs (within the limited range of sensitivities covered by the AOGCMs) or of a bias in the predicted sensitivity, subject to the caveat that we cannot, of necessity, sample errors that all models have in common using this approach.

A second verification of performance is given in Fig. 6, where we show the dependence of the prediction error and the predicted best guess and range of sensitivity as a function of the number of neurons. Increasing the number of neurons (i.e., increasing the number of degrees of freedom) decreases the error in the predicted sensitivity for both the CPDN verification set (1000 runs, Fig. 6a) and for the 14 AOGCMS [GISS and the Canadian Centre for Climate Modelling and Analysis (CCCma) models excluded; Fig. 6b]. However, increasing the neuron number beyond about five only improves the performance for the CPDN data, but not for the other AOGCMs. We interpret this as the fact that the neural network starts to fit CPDN specific

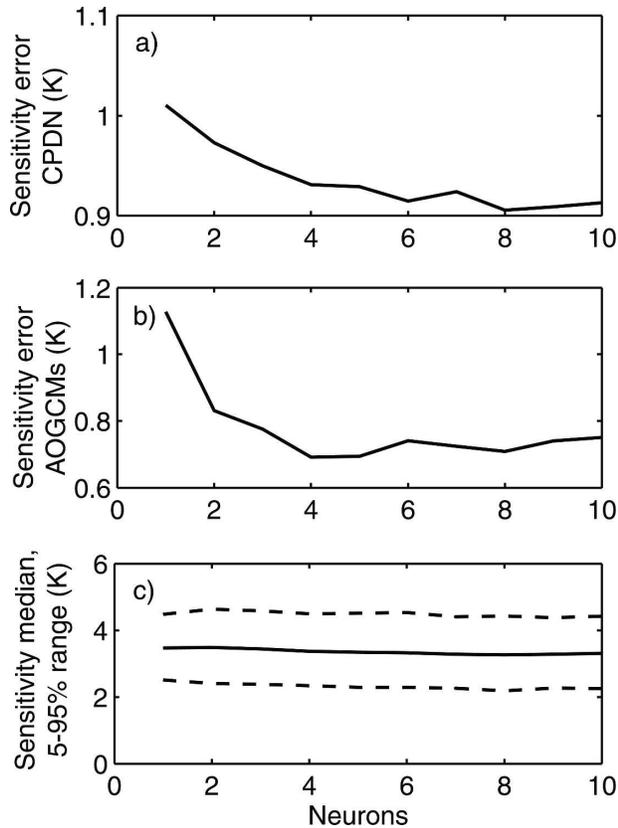


FIG. 6. (a) Error (1 std dev) of the predicted sensitivity on a subset of the CPDN data (not used for training), (b) error (1 std dev) of the predicted sensitivity for 14 AOGCMs (GISS and CGCM models excluded), and (c) mean (solid) and 5%–95% range for sensitivity (dashed) predicted from observed/reanalysis seasonality over land vs number of neurons in the neural network.

patterns beyond this point, which are not present in other models and probably also not in the reanalysis data. On the other hand, increasing the number of neurons only very weakly affects the best guess and uncertainty of the sensitivity predicted from reanalysis and observational data (Fig. 6c), so the choice of the neural network size is not critical here. Since it is desirable to have good performance, avoid overfitting, and to have few degrees of freedom for computational efficiency, we chose 10 neurons as the standard case for Fig. 4. The results are also insensitive to details in the design of the neural network and the number of simulations used for training in this standard setup.

The observational dataset HadCruT2v and the two reanalysis datasets, which are generated by assimilating observational data, should match the “true” climate evolution very closely, in particular for temperature, which is comparably simple to model and where many good observations exist. There are nevertheless system-

atic errors in these datasets that are not captured by the uncertainty estimate from interannual variability. Indeed, for the mean seasonal cycle, the spread between ERA-40, NCEP–NCAR, and HadCruT2v is substantially larger than the uncertainty due to natural variability in the mean of each dataset. The effect of taking into account the spread of the different datasets is discussed below. Another point to consider is the fact that the reanalysis datasets cover only a period of about 50 yr. It is therefore impossible to make any statement about the extreme tails of the distribution given the limited number of years to estimate interannual variability. We assume the distribution to be normal, and there is no evidence from the data against that. To test the effect of a particular realization of 50 yr of variability on the predicted sensitivity range, sensitivity was calculated for different ensemble members of the same AOGCM and the 1950–2000 all-forcing simulations. In most cases, differences in the 5% and 95% levels for sensitivity were very small, but it must be stated clearly here that it is fundamentally impossible to exclude with certainty the possibility of climate sensitivity being outside any range. It is equally impossible to give a 99% confidence level for it, given the limited information on the tails of the prior for our climatology (Frame et al. 2005). In other words, even if the solid PDF in Fig. 4 suggests essentially zero probability for sensitivities above 8 K, that conclusion is not justified from that method, and probably also not from any similar method. A value for climate sensitivity of 11 K as found by Stainforth et al. (2005) can still not be excluded. The correct statement is that there is a 1 in 20 chance for sensitivity being above about a certain level, with very little information above that level. But this itself is a vital step to quantify uncertainties in model projections.

An additional problem is how to account for structural uncertainties. In a similar study, Piani et al. (2005) found that just accounting for internal climate variability leads to very tight constraints on sensitivity but, more importantly, a clear inconsistency between the residual and the noise model. Given that for the amplitude of the seasonal cycle in most regions, the standard deviation across the three datasets is larger than the uncertainty in each dataset due to internal variability, we use the sum of the two as a more reasonable uncertainty estimate of the seasonal cycle. This accounts in a crude way for part of the structural uncertainties and therefore provides minimum estimate of the contribution of structural uncertainty. In a sensitivity test where we double the combined uncertainty (interannual variability plus spread across datasets) again, the peak of the PDF remains similar and the lower bound changes only slightly. However, the 95% level is more sensitive

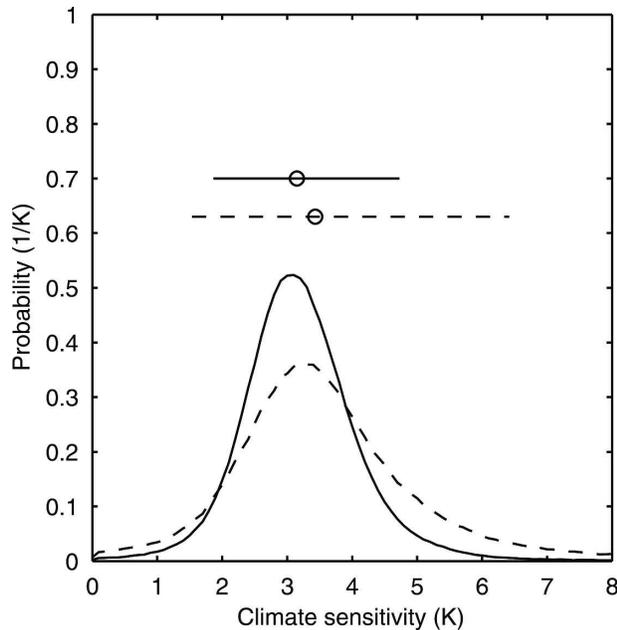


FIG. 7. PDFs, medians (circles), and 5%–95% ranges (horizontal lines) for climate sensitivity. The solid lines denote the case where the uncertainty in the observed seasonal cycle, in addition to natural variability, takes into account the spread across the ERA-40, NCEP–NCAR, and HadCruT2v datasets, and provides a minimum estimate of structural uncertainty to the PDF of climate sensitivity. Doubling this combined uncertainty again (dashed) indicates that the 95% level is most sensitive as to how structural uncertainties are treated. Medians are 3.1 and 3.4 K and the 5%–95% ranges are 1.9–4.7 and 1.5–6.4 K for the solid and dashed cases, respectively.

and the PDF becomes increasingly skewed, indicating that the upper bound on climate sensitivity is more difficult to constrain. Figure 7 shows the medians, 5%–95% ranges, and PDFs of climate sensitivity for the two cases.

We conclude that structural uncertainty is an important contribution to the total uncertainty here, but has received little attention in these types of studies so far. Even if there is little evidence from the sensitivities derived for other AOGCMs, we conclude from the arguments above that when trying to account for structural uncertainties in the datasets and the model used, the result in Fig. 4 does not capture the full uncertainty range, and that a more conservative estimate as in Fig. 7 would place an upper bound (95%) for climate sensitivity between 5 and 6.5 K.

An additional problem not considered is that there is a systematic relationship between climate sensitivity and the top-of-atmosphere energy imbalance in the climateprediction.net ensembles: these being slab models, there is no constraint that integrated fluxes at the surface and top of atmosphere sum to zero. Since many

observable quantities, including the seasonal cycle, would likely be affected by an overall atmospheric energy imbalance for reasons that have nothing to do with atmospheric feedbacks, this may have implications for the approach to constraining sensitivity that we propose here, which we will explore in subsequent work. Fortunately, the climateprediction.net ensemble is growing rapidly such that it will soon be possible to restrict attention to a subset of model versions that are approximately in energy balance, avoiding this potential problem.

5. Conclusions

The amplitude of the seasonal cycle in temperature provides a strong constraint on climate sensitivity. Using a neural network and climatological data, and subject to the assumptions discussed above, we find a median and 5%–95% range for climate sensitivity of 3.3 and 2.2–4.4 K when treating three observational and reanalysis datasets as equally plausible. This however ignores structural uncertainty and thus does not capture the full uncertainty range. Including a simple representation of structural uncertainty by increasing the uncertainty of the observations widens the uncertainty range. While the lower bound (5%) is relatively robust at 1.5–2 K, the upper one (95%) is sensitive to the assumptions made and varies between 5 and 6.5 K.

Unlike some studies based on the transient warming of the twentieth century, which show a large probability for high values of sensitivity, this method suggests an upper bound (95% level) on climate sensitivity at about 6.5 K for most cases considered, and above about 5 K when using an optimal combination of regions and degrees of freedom and making our most optimistic assumptions about the origins of model–data inconsistency. The 5%–95% confidence range is in broad agreement with the range of climate sensitivities spanned by the AOGCM models currently in use, but with a somewhat longer tail at high values that is not sampled by any AOGCM used for the upcoming IPCC report. Equally important as the constraints placed on the upper bound, this method indicates that climate sensitivity is very unlikely below 1.5–2 K, independent of all assumptions. This supports the view that the net feedbacks are substantially positive and it provides a lower bound on the climate change we have to expect in the future. While not allowing to completely exclude very high values of climate sensitivity, this approach allows one to attach probabilities to any value of sensitivity within the 5%–95% interval, and thus provides a basis for probabilistic projections based on large climate model ensembles.

Acknowledgments. We thank all the individuals who have contributed their computer time to make this project happen; Climate and Environmental Physics at the University of Bern for hosting and administrating a CPDN server; and Claudio Piani, Dave Frame, Ben Sanderson, Jacqueline Flückiger, Doug Nychka, Reinhard Furrer, Malte Meinshausen, and Thomas Stocker for stimulating discussions. Author RK was supported by the Swiss National Science Foundation. This study was supported in part by the Office of Biological and Environmental Research, U.S. Department of Energy, as part of its Climate Change Prediction Program. This work was also supported in part by the Weather and Climate Impact Assessment Initiative at the National Center for Atmospheric Research. We acknowledge the international modeling groups for providing their data for analysis: the Program for Climate Model Diagnosis and Intercomparison (PCMDI) for collecting and archiving the model data, the JSC/CLIVAR Working Group on Coupled Modelling (WGCM) and their Coupled Model Intercomparison Project (CMIP) and Climate Simulation Panel for organizing the model data analysis activity, and the IPCC WG1 TSU for technical support. The IPCC Data Archive at Lawrence Livermore National Laboratory is supported by the Office of Science, U.S. Department of Energy.

REFERENCES

- Andronova, N. G., and M. E. Schlesinger, 2001: Objective estimation of the probability density function for climate sensitivity. *J. Geophys. Res.*, **106** (D19), 22 605–22 611.
- Annan, J. D., J. C. Hargreaves, R. Ohgaito, A. Abe-Ouchi, and S. Emori, 2005: Efficiently constraining climate sensitivity with paleoclimate simulations. *SOLA*, **1**, 181–184.
- Braganza, K., D. J. Karoly, A. C. Hirst, M. E. Mann, P. Stott, R. J. Stouffer, and S. F. B. Tett, 2003: Simple indices of global climate variability and change: Part I—Variability and correlation structure. *Climate Dyn.*, **20**, 491–502.
- Charney, J. G., 1979: *Carbon Dioxide and Climate: A Scientific Assessment*. National Academy of Science, 22 pp.
- 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.
- , —, and —, 2006: Estimated PDFs of climate system properties including natural and anthropogenic forcings. *Geophys. Res. Lett.*, **33**, L01705, doi:10.1029/2005GL023977.
- Frame, D. J., 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, doi:10.1029/2004GL022241.
- Giorgi, F., and R. Francisco, 2000: Uncertainties in regional climate change predictions. *Climate Dyn.*, **16**, 169–182.
- Gregory, J. M., R. J. Stouffer, S. C. B. Raper, P. A. Stott, and N. A. Rayner, 2002: An observationally based estimate on the climate sensitivity. *J. Climate*, **15**, 3117–3121.
- Hagan, M. T., and M. Menhaj, 1994: Training feedforward networks with the Marquart algorithm. *IEEE Trans. Neural Networks*, **5**, 989–993.
- , H. B. Demuth, and M. H. Beale, 1996: *Neural Network Design*. PWS Publishing, 736 pp.
- Hall, A., and X. Qu, 2006: Using the current seasonal cycle to constrain snow albedo feedback in future climate change. *Geophys. Res. Lett.*, **33**, L03502, doi:10.1029/2005GL025127.
- 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.
- Hoffert, M. I., and C. Covey, 1992: Deriving global climate sensitivity from paleoclimate reconstructions. *Nature*, **360**, 573–576.
- Houghton, J. T., L. G. Meira Filho, B. A. Callander, N. Harris, A. Kattenberg, and K. Maskell, Eds., 1996: *Climate Change 1995: The Science of Climate Change*. Cambridge University Press, 572 pp.
- , Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, and D. Xiaosu, Eds., 2001: *Climate Change 2001: The Scientific Basis*. Cambridge University Press, 944 pp.
- Jones, P. D., and A. Moberg, 2003: Hemispheric and large-scale surface air temperature variations: An extensive revision and an update to 2001. *J. Climate*, **16**, 206–223.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.
- 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.
- , —, —, and —, 2003: Probabilistic climate change projections using neural networks. *Climate Dyn.*, **21**, 257–272.
- Lea, D., 2004: The 100 000-yr cycle in tropical SST, greenhouse forcing, and climate sensitivity. *J. Climate*, **17**, 2170–2179.
- Murphy, J. M., D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, and D. A. Stainforth, 2004: Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature*, **430**, 768–772.
- Myhre, G., E. J. Highwood, and K. P. Shine, 1998: New estimates of radiative forcing due to well mixed greenhouse gases. *Geophys. Res. Lett.*, **25**, 2715–2718.
- Piani, C., D. J. Frame, D. A. Stainforth, and M. R. Allen, 2005: Constraints on climate change from a multi-thousand member ensemble of simulations. *Geophys. Res. Lett.*, **32**, L23825, doi:10.1029/2005GL024452.
- Pope, V. D., M. L. Gallani, P. R. Rowntree, and R. A. Stratton, 2000: The impact of new physical parametrizations in the Hadley Centre climate model: HadAM3. *Climate Dyn.*, **16** (2–3), 123–146.
- Schneider von Deimling, T., H. Held, A. Ganopolski, and S. Rahmstorf, 2006: Climate sensitivity estimated from ensemble simulations of glacial climate. *Climate Dyn.*, **27**, 149–163.
- Simmons, A. J., and J. K. Gibson, 2000: The ERA-40 project plan. ERA-40 Project Rep. Series 1, European Centre for Medium-Range Weather Forecasting, Reading, United Kingdom, 63 pp.
- Stainforth, D. A., and Coauthors, 2005: Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature*, **433**, 403–406.
- Stott, P. A., and J. A. Kettleborough, 2002: Origins and estimates of uncertainty in predictions of twenty-first century temperature rise. *Nature*, **416**, 723–726.
- Wigley, T. M. L., and S. C. B. Raper, 2001: Interpretation of high projections for global-mean warming. *Science*, **293**, 451–454.