Broad range of 2050 warming from an observationally constrained large climate model ensemble

Daniel J. Rowlands[1,2,3]★, David J. Frame[1,2,4], Duncan Ackerley[5,6], Tolu Aina[1,7], Ben B. B. Booth[8], Carl Christensen[1], Matthew Collins[8,9], Nicholas Faull[1], Chris E. Forest[10], Benjamin S. Grandey[1], Edward Gryspeerdt[1], Eleanor J. Highwood[6], William J. Ingram[1,8], Sylvia Knight[11], Ana Lopez[3], Neil Massey[1,4], Frances McNamara[12], Nicolai Meinshausen[13], Claudio Piani[14,15], Suzanne M. Rosier[1], Benjamin M. Sanderson[16], Leonard A. Smith[3,17], Dáithí A. Stone[18], Milo Thurston[7], Kuniko Yamazaki[1], Y. Hiro Yamazaki[1,19] and Myles R. Allen[1,2,4]

Incomplete understanding of three key properties of the climate system—equilibrium climate sensitivity, rate of ocean heat uptake and historical aerosol forcing—and their underlying physical processes lead to uncertainties in our assessment of the global-mean temperature evolution in the twenty-first century[1,2]. Explorations of these uncertainties have so far relied on scaling approaches[3,4], large ensembles of simplified climate models[1,2], or small ensembles of complex coupled atmosphere–ocean general circulation models[5,6], which under-represent uncertainties in key climate system properties derived from independent sources[7-9]. Here we present results from a multi-thousand-member perturbed-physics ensemble of transient coupled atmosphere–ocean general circulation model simulations. We find that model versions that reproduce observed surface temperature changes over the past 50 years show global-mean temperature increases of 1.4–3 K by 2050, relative to 1961–1990, under a mid-range forcing scenario. This range of warming is broadly consistent with the expert assessment provided by the Intergovernmental Panel on Climate Change Fourth Assessment Report[10], but extends towards larger warming than observed in ensembles-of-opportunity[5] typically used for climate impact assessments. From our simulations, we conclude that warming by the middle of the twenty-first century that is stronger than earlier estimates is consistent with recent observed temperature changes and a mid-range 'no mitigation' scenario for greenhouse-gas emissions.

In the latest generation of coupled atmosphere–ocean general circulation models (AOGCMs) contributing to the Coupled Model Intercomparison Project phase 3 (CMIP-3), uncertainties in key properties controlling the twenty-first century response to sustained anthropogenic greenhouse-gas forcing were not fully sampled, partially owing to a correlation between climate sensitivity and aerosol forcing[7,8], a tendency to overestimate ocean heat uptake[11,12] and compensation between short-wave and long-wave feedbacks[9]. This complicates the interpretation of the ensemble spread as a direct uncertainty estimate, a point reflected in the fact that the 'likely' (>66% probability) uncertainty range on the transient response was explicitly subjectively assessed as −40% to +60%

of the CMIP-3 ensemble mean for global-mean temperature in 2100, in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4). The IPCC expert range was supported by a range of sources[10], including studies using pattern scaling[3,4], ensembles of intermediate-complexity models[1,2] and estimates of the strength of carbon-cycle feedbacks[13]. From this evidence it is clear that the CMIP-3 ensemble, which represents a valuable expression of plausible responses consistent with our current ability to explore model structural uncertainties, fails to reflect the full range of uncertainties indicated by expert opinion and other methods.

In the absence of uncertainty guidance or indicators at regional scales, studies have relied on the CMIP-3 ensemble spread as a proxy for response uncertainty[14], or statistical post-processing to correct and inflate uncertainty estimates[15], at the risk of violating the physical constraints provided by dynamical AOGCM simulations, especially when extrapolating beyond the range of behaviour in the raw ensemble.

Perturbed-physics ensembles[6,16,17] offer a systematic approach to quantify uncertainty in models of the climate system response to external forcing. Here we investigate uncertainties in the twenty-first century transient response in a multi-thousand-member ensemble of transient AOGCM simulations from 1920 to 2080 using HadCM3L, a version of the UK Met Office Unified Model, as part of the climate*prediction*.net British Broadcasting Corporation (BBC) climate change experiment (CCE). We generate ensemble members by perturbing the physics in the atmosphere, ocean and sulphur cycle components, with transient simulations driven by a set of natural forcing scenarios and the SRES A1B emissions scenario[18], and also control simulations to account for unforced model drifts (Methods and Supplementary Fig. S1).

Figure 1 shows the evolution of global-mean surface temperatures in the ensemble (relative to 1961–1990), each coloured by the goodness-of-fit to observations of recent surface temperature changes, as detailed below. The raw ensemble range (1.1–4.2 K around 2050), primarily driven by uncertainties in climate sensitivity (Supplementary Information), is potentially misleading because many ensemble members have an unrealistic response to the forcing over the past 50 years. We compare model-simulated spatial aver-

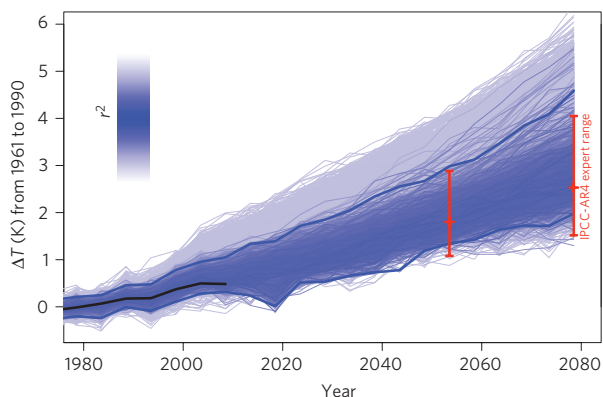A full list of affiliations appears at the end of the paper.

Figure 1 | Evolution of uncertainties in reconstructed global-mean temperature projections under SRES A1B in the HadCM3L ensemble. Blue colouring indicates goodness-of-fit between observations and ensemble members, plotted in order of increasing agreement (light to dark blue). Black line, the evolution of observations, and thick blue lines the 'likely' range (66% confidence interval) from the ensemble. Red bars show the IPCC-AR4 expert 'likely' range around 2050 and 2080. All temperatures are relative to the corresponding 1961–1990 mean. For consistency and to account for the observational mask, global-means are reconstructed from Giorgi + ocean region averages (0.2 K less on average).

ages (Giorgi regions and ocean basins, Supplementary Table S3 and Fig. S7) of five-year mean surface (1.5 m) temperature changes over 1961–2010 with observations[19], all expressed as anomalies from the respective 1961 to 1990 mean. We test model versions against regional temperature changes over the past 50 years because they have been shown to correlate well with forecast future warming[3], whereas mean temperatures do not[20]. We constrain the model base climatology by filtering the ensemble to retain only model versions requiring a global annual mean flux adjustment in the range ±5 W m⁻², comparable to estimates of observational uncertainty in top-of-atmosphere fluxes[6].

Assessing goodness-of-fit, which represents a limited expression of model error, requires a measure of the expected error between model simulations and observations due to sampling uncertainty, assuming it is primarily from internally-generated climate variability. We estimated variability using segments of long pre-industrial control simulations from CMIP-3, filtered to retain spatial scales on which AOGCM-based estimates of variability are reliable (Supplementary Fig. S8). We focus on the range of projections provided by model versions that satisfy a given goodness-of-fit threshold, rather than explicitly weighting model versions, given the sensitivity of results to noise in individual simulations[21] and parameter sampling design[22].

Figure 2a shows that without a goodness-of-fit threshold, hindcasts of 2001–2010 global-mean warming relative to 1961–1990 show a wide range from 0 to 1.5 K. We define a 'likely' range (66% confidence interval) by considering the range from ensemble members with model error (y-axis) lower than the 66th percentile of the distribution of model error arising from estimates of internal variability alone (black crosses), giving a range of 0.3–0.9 K. This is the range of warming to date that we estimate might have occurred at this confidence level given the evidence of our ensemble and estimates of modelled internal climate variability from CMIP-3. The observed warming of 0.5 K is close to our best-fit model version (not identical, as we use more than just global-mean trend information in our goodness-of-fit measure), and 0.1 K below the centre of our uncertainty range. This is consistent with temperatures over 2001–2010 being slightly depressed by a combination of internal variability[23] and two factors not sampled in our ensemble: stratospheric water vapour decreases[24] and an unusually low solar minimum[25]. Note that the grey bar represents observational uncertainty in the warming that actually occurred, while our constrained ensemble range represents the warming that might have occurred over this period given internal variability and response uncertainty.

On the assumption that models that simulate past warming realistically are our best candidates for making estimates of the future, we find a 'likely' range of 1.4–3 K for warming around 2050 under the SRES A1B scenario (Fig. 2b). No ensemble members warm by less than 1 K by 2050 under this scenario, despite the large size of the ensemble and allowance for natural forcing uncertainty: we allow explicitly for future volcanic activity and include a scenario
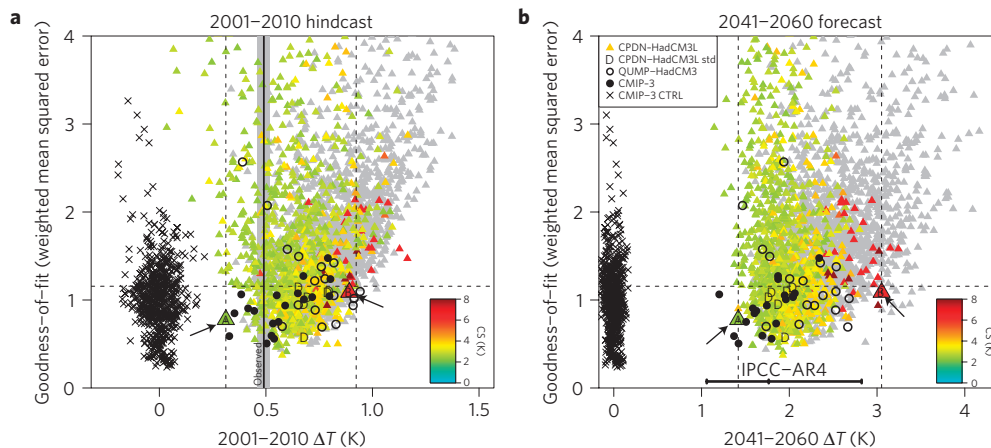

Figure 2 | Goodness-of-fit to recent temperature changes as a function of global-mean warming. a, 2001–2010 reconstructed hindcast; **b**, 2041–2060 forecast under SRES A1B for global-mean temperature, both as anomalies from 1961 to 1990. Coloured triangles, members of the HadCM3L perturbed-physics ensemble, with colours denoting the corresponding slab model estimated equilibrium climate sensitivity. D symbols, standard physics model configurations differing in natural forcing scenario and scaling on anthropogenic sulphate emissions. Black crosses, realizations of model error and corresponding temperature changes arising from simulations of internal variability, with the horizontal line denoting the 66th percentile of the error distribution. Vertical dotted lines, the range of the HadCM3L ensemble with errors lower than this percentile corresponding to a 'likely' range (66% confidence interval). Grey triangles, simulations with global annual mean flux adjustments outside ±5 W m⁻². Black vertical bar and grey band in **a**, observations and 'likely' range. Horizontal bar in **b**, the expert IPCC-AR4 'likely' range. Black filled circles CMIP-3 simulations, black open circles QUMP HadCM3 simulations. Arrowed larger triangles refer to models highlighted in Fig. 3.
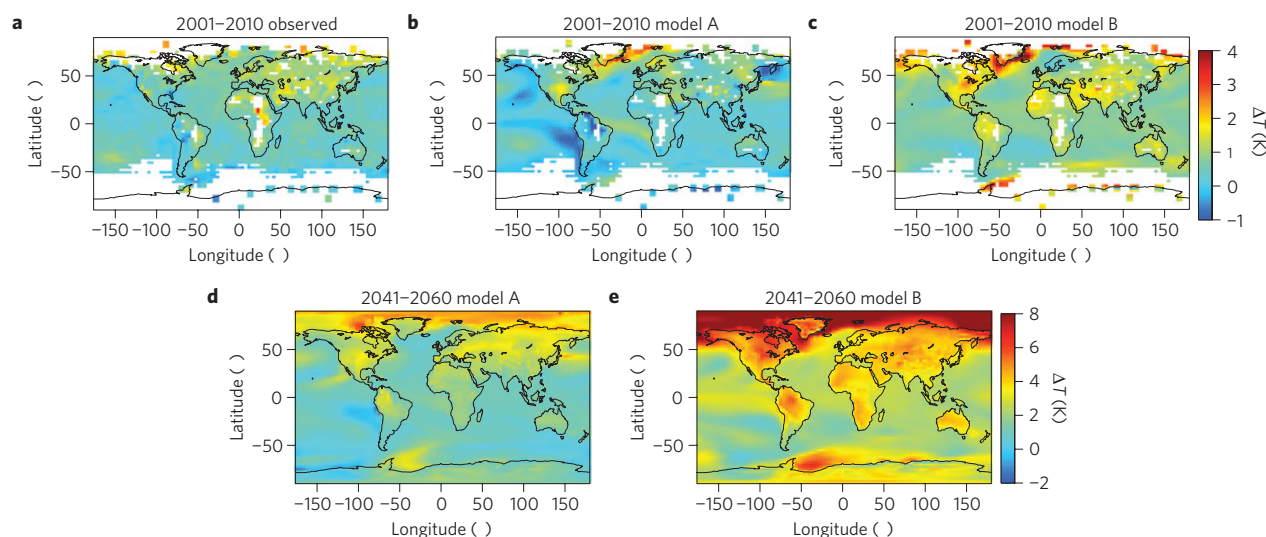
Figure 3 | Surface temperature anomaly fields relative to 1961–1990 for 2001–2010 hindcast and 2041–2060 forecast for a low-response ensemble member, A ($\Delta T_{2050} = 1.4$ K), and high-response ensemble member, B ($\Delta T_{2050} = 3$ K) labelled in Fig. 2. a, Observed 2001–2010 anomaly; **b,d** model A anomaly for 2001–2010 and 2041–2060; **c,e** model B anomaly. White regions in **a–c** indicate missing data, defined as > 40% of yearly data missing over 1961–1990 or 2001–2010. The same mask is applied in **b** and **c**. Note the factor of two difference in colour-scale between **a–c** and **d,e**.

in which solar activity falls back to 1900 levels. This finding is compatible with energy balance calculations[26], given the level of greenhouse-gas forcing by 2050 and the lower limit of climate sensitivity explored in the ensemble at approximately 2 K, close to the lower end of the range of sensitivities considered 'likely' by the IPCC-AR4 (ref. 10).

The lower end of our 'likely' range for 2050 warming at 1.4 K is consistent with the lowest responses in the CMIP-3 ensemble (filled circles Fig. 2b), lower than the lowest realistic (on this measure) members of the UK Met-Office QUMP HadCM3 perturbed-physics ensemble[6] (open circles Fig. 2b), and higher than IPCC expert lower bound[10] (the CMIP-3 ensemble-mean minus 40%). This is contingent evidence that the real-world response is likely to be at least as large as the lowest responses in the CMIP-3 ensemble, and that the IPCC-AR4 expert estimate of the lower bound was probably over-conservative. This comparison with the IPCC expert assessment is valid under the assumption of constant fractional uncertainty in the twenty-first century response[3,8], given that the IPCC expert estimate was stated for 2100.

At about 3 K, the upper end of our uncertainty range for 2050 warming is consistent with both the highest responses in the QUMP ensemble and the IPCC upper estimate of the CMIP-3 ensemble-mean plus 60% (ref. 10), but substantially higher than the highest responses of the CMIP-3 ensemble members that are generally used for impact assessment (one model did give a higher response, but was omitted in headline uncertainty ranges because of concerns about its stability). Thus uncertainty estimates based solely on ensembles-of-opportunity or small perturbed-physics ensembles are underestimated compared with independent studies[2,3]. We are reluctant to quote a more precise upper bound because of the small number of model versions in this region and the fact that goodness-of-fit does not deteriorate as rapidly as it does at the lower bound, possibly because of the inclusion of natural forcing uncertainty: we can, however, conclude that warming substantially greater than 3 K by 2050 is unlikely unless forcing is substantially higher than the A1B scenario[27]. The higher upper bound compared to CMIP-3 is mostly due to our inclusion of a wider range of climate sensitivities but also partly to our wider range of natural forcing scenarios (Supplementary Figs S1 and S4). Towards the end of the century, we observe a similar relationship with the IPCC expert estimate (red bar, Fig. 1), although by that time

the uncertainty could be larger if carbon-cycle feedbacks were included in our ensemble[13].

To the extent that policy makers require 'a range of plausible representations of future climate'[28], providing uncertainty guidance in this way can have an important role to play. Further observational constraints may reduce uncertainty further, particularly those relating to forced responses such as the seasonal cycle[29], although the use of seasonally varying flux adjustments here may distort any relationship. We find little sensitivity in our results to varying the flux adjustment threshold and removing this constraint entirely adds approximately 0.4 K to the upper bound in 2050 through admitting a number of high climate sensitivity model versions (Supplementary Fig. S10). Conversely, we are likely to have undersampled uncertainty in ocean heat uptake arising from ocean physics through perturbing only a single, coarse-resolution, ocean model structure[6]: more generally, sampling structural uncertainty might allow for the impact of further observational constraints such as ocean heat content changes.

Perhaps unexpectedly, we observe little relationship between climate sensitivity and aerosol forcing (as measured through the sulphate burden) in the constrained ensemble (Supplementary Fig. S2). We attribute this to the choice of the 1961–1990 reference period for the transient-control anomaly, which removes much of the spread across the ensemble arising from aerosol forcing uncertainty (Supplementary Figs S3 and S4). Filtering the ensemble based on a comparison of the modelled values of the mean transient-control anomaly over 1961–1990 to the observed warming between 1891–1910 (representative of the control simulations in the ensemble) and 1961–1990, reduces our upper bound to approximately 2.8 K. The design of the experiment, whereby simulations were launched in 1920, precludes us from applying this as a formal constraint given the difficulty of comparing the control simulation like-for-like to any period in the past, in addition to the paucity of observations at the start of the twentieth century.

Unlike uncertainty estimates based on intermediate-complexity models[11], pattern-scaling[4] or statistical emulation[15], every member of the BBC CCE is consistent with physical constraints as expressed in the HadCM3L AOGCM, ensuring physical coherence of results for investigating joint uncertainties. However, the BBC CCE clearly does not sample model errors common to all of the

current generation of AOGCMs that arise from limited process understanding and computational resources.

Figure 3 shows surface warming in a low-response (Model A, global $\Delta T_{2050} = 1.4\,\text{K}$) and high-response (Model B, global $\Delta T_{2050} = 3\,\text{K}$) ensemble member. Model A and B show a larger difference in the contrast of Pacific equatorial warming—specifically the Niño 3.4 region—relative to the warming over the whole Pacific Ocean, when compared with the corresponding range observed in either the CMIP-3 or QUMP ensembles, providing evidence that perturbed-physics ensembles can sample spatial response uncertainty.

Uncertainty estimates for the transient response are conditioned on a given emissions scenario[10]. For the SRES A1B scenario, we have shown that a more complete sampling of uncertainty in key climate system properties and forcings produces a wider range of projections for the coming century consistent with recent surface temperature observations than in the CMIP-3 ensemble used for regional projections in IPCC-AR4, and similar to the IPCC authors' expert assessment of uncertainty in the global response. Reliance on the spread of responses in an ensemble-of-opportunity can underestimate uncertainties, particularly at the upper end of the range for twenty-first century warming. Our ensemble provides a set of physically coherent simulations consistent with recent observed warming, giving plausible worlds beyond the range generated by ensembles-of-opportunity which can aid the development of robust climate adaptation policies.

Methods

Model simulations. HadCM3L consists of a 3.75° longitude by 2.5° latitude atmosphere with interactive sulphur cycle coupled to a dynamical ocean of the same resolution[30]. Model physics parameters are perturbed through expert elicitation[16], and informed for atmospheric and sulphur cycle physics perturbations by results from the climate*prediction*.net slab model experiment[30], choosing between two and four values for each parameter (Supplementary Tables S1 and S2). Atmospheric configurations are initially chosen to span a wide range of equilibrium climate sensitivities (2–9 K, estimated from the slab ocean model experiments) whilst still retaining an acceptable climatology, measured through the top-of-atmosphere flux imbalance relative to the standard physics settings ($\pm 10\,\text{W m}^{-2}$).

Flux adjustments are calculated for 10 ocean configurations through a 200-year spin-up coupled to the standard atmosphere, and for each of 153 perturbed atmospheres[30], producing 1,530 possible model versions. For each model version the flux adjustments were applied in two initial condition ensembles of 160-year simulations: (1) control simulations with constant forcing (representative of 1880–1920 mean conditions) to check and allow for unforced drifts and (2) transient simulations from 1920 to 2080 forced with changes in greenhouse gases and a set of sulphate emissions under the SRES A1B emissions scenario[18], together with a set of solar and volcanic forcing scenarios (Supplementary Fig. S1).

In total 9,745 complete simulations were returned from the climate*prediction*.net participants. Given bandwidth and storage constraints in the distributed computing environment, we restrict our analysis to surface temperature data focussing on 22 Giorgi land regions and 6 major ocean basins for our comparison with observations (Supplementary Table S3).

Data preparation. Of the 9,745 complete simulations there are 1,656 controls and 8,089 transients. Model versions with absolute global-mean drifts in the control climate larger than 0.4 K/century are flagged, indicating the flux adjustment has not eliminated unforced drifts. Transient simulations are matched based on their parameters and natural forcing scenario. Initial condition ensemble averages are taken where possible to reduce noise in the model simulations. Controls are prepared identically, and matched to corresponding transients through the model parameters, giving a total of 2,752 distinct transient-control pairs. The 2,752 transient-control pairs contain 809 of the original 1,530 possible model versions. Each transient-control pair is expressed as an anomaly from the 1961 to 1990 mean in each region. Observations, from HadCRUT3 (ref. 19), AOGCM simulations under the A1B scenario and CMIP-3 pre-industrial control segments are prepared identically (Supplementary Table S4). Finally, all data is temporally averaged to 5-year mean resolution to reduce the impact of internal variability. For simplicity, coverage is assumed complete within Giorgi regions in this analysis of the model output: this introduces only small errors because the regions used have a high observational coverage (around 95% for each 5-year period) over the 1961–2010 period considered (Fig. 3a).

Goodness-of-fit calculation. We calculate a goodness-of-fit statistic based on the spatio-temporal pattern of surface temperature from 1961 to 2010 as

$$r_\theta^2 = \left(\mathbf{y} - \mathbf{x}_\theta\right)^T \mathbf{C}_N^{-1}\left(\mathbf{y} - \mathbf{x}_\theta\right)$$

where $\mathbf{y}$ represents observations, $\mathbf{x}_\theta$ a transient-control pair of simulations corresponding to parameters θ, and $\mathbf{C}_N$ a covariance matrix which weights errors corresponding to the estimated variability in components of $\mathbf{y}$ and $\mathbf{x}_\theta$ arising from internal climate variability. Observations cannot accurately be used to estimate $\mathbf{C}_N$ without simplifying assumptions, and so standard practice is to use segments of pre-industrial control simulations[3]. We use pre-industrial control simulations from all available CMIP-3 models to account for variability in $\mathbf{y}$, and a 1,000 year HadCM3 control run to characterize variability in $\mathbf{x}_\theta$. We find little sensitivity in the results to scaling the variability associated with $\mathbf{y}$ over a wide range (Supplementary Fig. S12).

Estimates of variability from AOGCMs are most reliable on large spatial scales, so we focus on the leading Empirical Orthogonal Functions (EOFs) of the HadCM3L ensemble over 1961–2010, the first three of which explain over 90% of the spatial variance across the ensemble (Supplementary Fig. S6). The exact choice of truncation does not significantly impact results when using a regularized covariance estimate (Supplementary Equation S19), and using a separate physically-based dimension reduction technique does not change our conclusions (Supplementary Fig. S11).

For a given confidence level, we compare r_θ^2 with the corresponding percentile of the distribution of r^2 arising from estimates of internal variability alone using the pre-industrial control segments. A schematic of the analysis is shown in Supplementary Fig. S5. We use an independent set of control segments to $\mathbf{C}_N$ to remove the small sample size bias[3]. We test the null hypothesis that the model and observations come from the same distribution and reject the model simulation if r_θ^2 is too large. In Fig. 2 we show goodness-of-fit as a weighted mean squared error by normalizing r_θ^2 by the number of degrees of freedom in $\mathbf{y}$ and $\mathbf{x}_\theta$. For reference, model simulations must explain at least 50% of the variance in filtered surface temperature observations to pass the r^2 test.

References

1. Forest, C. E., Stone, P. H., Sokolov, A. P., Allen, M. R. & Webster, M. D. Quantifying uncertainties in climate system properties with the use of recent climate observations. *Science* **295,** 113–116 (2002).
2. Knutti, R., Stocker, T. F., Fortunat, J. & Plattner, G. K. Constraints on radiative forcing and future climate change from observations and climate model ensembles. *Nature* **416,** 719–723 (2002).
3. Stott, P. A. *et al.* Observational constraints on past attributable warming and predictions of future global warming. *J. Clim.* **19,** 3055–3069 (2006).
4. Harris, G. R. *et al.* Frequency distributions of transient regional climate change from perturbed-physics ensembles of general circulation model simulations. *Clim. Dynam.* **27,** 357–375 (2006).
5. Meehl, G. A. *et al.* The WCRP CMIP3 multimodel dataset: A new era in climate change research. *Bull. Am. Meteorol. Soc.* **88,** 1383–1394 (2007).
6. Collins, M. *et al.* Climate model errors, feedbacks and forcings: A comparison of perturbed-physics and multi-model ensembles. *Clim. Dynam.* **36,** 1737–1766 (2010).
7. Kiehl, J. Twentieth century climate model response and climate sensitivity. *Geophys. Res. Lett.* **34,** L22710 (2007).
8. Knutti, R. Why are climate models reproducing the observed global surface warming so well? *Geophys. Res. Lett.* **35,** L18704 (2008).
9. Huybers, P. Compensation between model feedbacks and curtailment of climate sensitivity. *J. Clim.* **23,** 3009–3018 (2010).
10. Knutti, R. *et al.* A review of uncertainties in global temperature projections over the twenty-first century. *J. Clim.* **21,** 2651–2663 (2008).
11. Forest, C. E., Stone, P. H. & Sokolov, A. P. Constraining climate model parameters from observed 20th century changes. *Tellus A* **60,** 911–920 (2008).
12. Boé, J., Hall, A. & Qu, X. Deep ocean heat uptake as a major source of spread in transient climate change simulations. *Geophys. Res. Lett.* **36,** L22701 (2009).
13. Friedlingstein, P. *et al.* Climate-carbon cycle feedback analysis: Results from the C4MIP model intercomparison. *J. Clim.* **19,** 3337–3353 (2006).
14. Milly, P. C. D., Dunne, K. A. & Vecchia, V. Global pattern of trends in stream flow and water availability in a changing climate. *Nature* **428,** 347–350 (2005).
15. Tebaldi, C. & Sansó, B. Joint projections of temperature and precipitation change from multiple climate models: A hierarchical Bayesian approach. *J. R. Stat. Soc. A* **172,** 83–106 (2009).
16. Murphy, J. M. *et al.* Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature* **430,** 768–772 (2004).
17. Jackson, C. S., Sen, M. K., Huerta, G., Deng, Y. & Bowman, K. P. Error reduction and convergence in climate prediction. *J. Clim.* **21,** 6698–6709 (2008).

18. Nakicenovic, N. & Swart, R. *Special Report on Emissions Scenarios* (Cambridge Univ. Press, 2000).

19. Brohan, P., Kennedy, J. J., Harris, I., Tett, S. F. B. & Jones, P. D. Uncertainty estimates in regional and global observed temperature changes: A new data set from 1950. *J. Geophys. Res.* **111**, D12106 (2006).

20. Knutti, R., Furrer, R., Tebaldi, C., Cermak, J. & Meehl, G. A. Challenges in combining projections from multiple climate models. *J. Clim.* **23**, 2739–2758 (2010).

21. Weigel, A. P., Knutti, R., Liniger, M. & Appenzeller, C. Risks of model weighting in multimodel climate projections. *J. Clim.* **23**, 4175–4191 (2010).

22. Frame, D. J. *et al.* Constraining climate forecasts: The role of prior assumptions. *Geophys. Res. Lett.* **32**, L09702 (2005).

23. Easterling, D. R. & Wehner, M. F. Is the climate warming or cooling? *Geophys. Res. Lett.* **36**, L08706 (2009).

24. Solomon, S. *et al.* Contributions of stratospheric water vapor to decadal changes in the rate of global warming. *Science* **327**, 1219–1223 (2010).

25. Lockwood, M. Solar change and climate: An update in the light of the current exceptional solar minimum. *Phil. Trans. R. Soc. Lond. A* **466**, 303–329 (2010).

26. Stone, D. A. & Allen, M. R. Attribution of global surface warming without dynamical models. *Geophys. Res. Lett.* **32**, L18711 (2005).

27. Betts, R. A. *et al.* When could global warming reach 4 °C. *Phil. Trans. R. Soc. Lond. A* **369**, 67–84 (2011).

28. Desaii, S., Hulme, M., Lempert, R. & Pielke, R. Jr Do we need better predictions to adapt to a changing climate? *Eos* **90**, 111–112 (2009).

29. Hall, A. & Qu, X. Using the current seasonal cycle to constrain snow albedo feedback in future climate change. *Geophys. Res. Lett.* **33**, L03502 (2006).

30. Frame, D. J. *et al.* The climate*prediction*.net BBC climate change experiment: Design of the coupled model ensemble. *Phil. Trans. R. Soc. Lond. A* **367**, 855–870 (2009).

Author contributions

All authors contributed to the design and implementation of the experiment. D.J.R. performed the analysis and wrote the paper, with significant contributions from D.J.F., M.R.A. and N.M. All authors commented on the paper.

Additional information

The authors declare no competing financial interests. Supplementary information accompanies this paper on www.nature.com/naturegeoscience. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.J.R.

[1]Atmospheric, Oceanic & Planetary Physics, Department of Physics, University of Oxford, Parks Road, Oxford OX1 3PU, UK, [2]School of Geography and the Environment, University of Oxford, South Parks Road, Oxford OX1 3QY, UK, [3]Centre for the Analysis of Time Series, London School of Economics, London WC2A 2AE, UK, [4]Smith School of Enterprise and the Environment, Hayes House, 75 George Street, Oxford OX1 2BQ, UK, [5]Monash Weather and Climate, Monash University, Clayton, Victoria 3800, Australia, [6]Department of Meteorology, University of Reading, Earley Gate, Reading, RG6 6BB, UK, [7]Oxford e-Research Centre, Keble Road, Oxford OX1 3QG, UK, [8]Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PU, UK, [9]College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, EX4 4QJ, UK, [10]Department of Meteorology, Earth and Environmental Systems Institute, Pennsylvania State University, University Park, Pennsylvania 16802, USA, [11]Royal Meteorological Society, Reading, RG1 7LL, UK, [12]BBC Science, BBC White City, 201 Wood Lane, London W12 7TS, UK, [13]Department of Statistics, University of Oxford, 1 South Parks Road, Oxford OX1 3TG, UK, [14]Abdus Salam International Center for Theoretical Physics, Trieste, Italy, [15]The American University of Paris, Paris, France, [16]National Center for Atmospheric Research, 1850 Table Mesa Dr, Boulder, Colorado 80305, USA, [17]Pembroke College, Oxford University of Oxford, Oxford OX1 1DW, UK, [18]Climate Systems Analysis Group, University of Cape Town, South Africa, [19]School of Geography, Politics and Sociology, Newcastle University, Newcastle on Tyne, NE1 7RU, UK.
*e-mail: rowlands@atm.ox.ac.uk.

Query 1:

Please note that the title has been changed according to style.

Query 2:

Please note that the first paragraph has been edited according to style.

Query 3:

Please provide postcode for affiliations 14,15.