nature climate change

LETTERS

https://doi.org/10.1038/s41558-018-0197-7

Corrected: Author Correction

Climate–carbon cycle uncertainties and the Paris Agreement

P. B. Holden[1*], N. R. Edwards[1,2], A. Ridgwell[3], R. D. Wilkinson[4], K. Fraedrich[5], F. Lunkeit[6], H. Pollitt[2,7], J.-F. Mercure[2,7,8], P. Salas[2], A. Lam[2,9], F. Knobloch[2,8], U. Chewpreecha[7] and J. E. Viñuales[2]

The Paris Agreement[1] aims to address the gap between existing climate policies and policies consistent with "holding the increase in global average temperature to well below 2 C". The feasibility of meeting the target has been questioned both in terms of the possible requirement for negative emissions[2] and ongoing debate on the sensitivity of the climate-carbon-cycle system[3]. Using a sequence of ensembles of a fully dynamic three-dimensional climate–carbon-cycle model, forced by emissions from an integrated assessment model of regional-level climate policy, economy, and technological transformation, we show that a reasonable interpretation of the Paris Agreement is still technically achievable. Specifically, limiting peak (decadal) warming to less than 1.7 °C, or end-of-century warming to less than 1.54 °C, occurs in 50% of our simulations in a policy scenario without net negative emissions or excessive stringency in any policy domain. We evaluate two mitigation scenarios, with 200 gigatonnes of carbon and 307 gigatonnes of carbon post-2017 emissions respectively, quantifying the spatio-temporal variability of warming, precipitation, ocean acidification and marine productivity. Under rapid decarbonization decadal variability dominates the mean response in critical regions, with significant implications for decision-making, demanding impact methodologies that address non-linear spatio-temporal responses. Ignoring carbon-cycle feedback uncertainties (which can explain 47% of peak warming uncertainty) becomes unreasonable under strong mitigation conditions.

A widely held misconception is that given the approximately 1 °C warming to date, and considering the committed warming (warming that will inevitably happen) concealed by ocean thermal inertia, the 1.5 °C target of the Paris Agreement[1] is already impossible. However, it is cumulative emissions that define peak warming[4]. When carbon emissions cease, terrestrial and marine sinks are projected to draw down atmospheric carbon dioxide (CO_2), approximately cancelling the lagging warming. Although the sign of this 'zero emissions commitment' is uncertain, its contribution can be neglected for low-CO_2 scenarios[5]. Therefore, at least when considering CO_2 emissions in isolation, keeping below 1.5 °C of warming will remain physically achievable until the point that it is reached. The physical achievability of the Paris target has been demonstrated in a complex carbon-cycle model with a simplified atmosphere[6] and updated recently using a simple carbon-cycle model forced by a modified RCP2.6 scenario[7] and by policy-driven scenarios that rely heavily on negative emissions technology[8]. Here, we demonstrate that the target is achievable using a fully dynamic three-dimensional climate–carbon-cycle model forced with emissions from a detailed set of sectorally and regionally specific mitigation policies without net negative emissions (see Methods).

We use the intermediate-complexity three-dimensional Earth system model PLASIM-GENIE[9], a model with ocean, atmosphere and carbon-cycle dynamics similar to those of full-complexity models, but with simpler parameterizations and lower spatial resolution. The model will not produce the full range of small-scale variability in high-complexity models, but it has the computational efficiency to allow a comprehensive treatment of uncertainties, taking into account, for instance, ongoing discussions on the state dependency of climate sensitivity[10,11] and ocean heat uptake efficacy[12]. We evaluate climate–carbon-cycle uncertainty using a 69-member history-matched[13] ensemble designed using 940 training simulations (see Methods). The ensemble climate sensitivity is 2.6 °C to 4.5 °C (90% confidence), which compares to 1.9 °C to 4.5 °C in the Coupled Model Intercomparison Project 5 (CMIP5)[14]. The transient climate response is 1.1 °C to 1.8 °C and 1.2 °C to 2.4 °C in CMIP5[14]. The ensemble ocean heat uptake (1965 to 2004) is 207–330 ZJ, 182–363 ZJ (1970 to 2010) in the Intergovernmental Panel on Climate Change (IPCC)[14].

We validate the history-matched ensemble in Table 1, by comparison with the CMIP5 multi-model ensembles forced by Representative Concentration Pathway (RCP) 2.6 (a mitigation scenario) and RCP8.5 (a 'business-as-usual' scenario)[15]. Under RCP8.5, the PLASIM-GENIE end-of-century CO_2 concentration, global warming and Atlantic Meridional Overturning Circulation (AMOC) strength[14,16] are remarkably consistent with the CMIP5 ensemble, illustrating that uncertainties in transient climate sensitivity, carbon-cycle sensitivity and AMOC stability capture the spread of high-complexity models. Mean surface pH is also well represented; the substantially lower uncertainty in CMIP5 pH[17] arises because these particular CMIP5 simulations were concentration-forced. Overstated impacts in marine productivity are apparent relative to CMIP5[17], but there is significant overlap in the highly uncertain distributions. Under RCP2.6 forcing, there is a less complete analysis of CMIP5 outputs. The PLASIM-GENIE ensemble understates the mean warming in RCP2.6 by 0.3 °C relative

[1]Environment, Earth and Ecosystem Sciences, The Open University, Milton Keynes, UK. [2]Cambridge Centre for Environment, Energy and Natural Resource Governance (C-EENRG), University of Cambridge, Cambridge, UK. [3]Department of Earth Sciences, University of California, Riverside, CA, USA. [4]School of Mathematics and Statistics, University of Sheffield, Sheffield, UK. [5]Max Planck Institute of Meteorology, KlimaCampus, Hamburg, Germany. [6]Meteorological Institute, University of Hamburg, Hamburg, Germany. [7]Cambridge Econometrics Ltd, Cambridge, UK. [8]Faculty of Science, Radboud University, Nijmegen, The Netherlands. [9]Department of Economics, Faculty of Social Science, University of Macao, Taipa, Macao.
*e-mail: philip.holden@open.ac.uk

Table 1 | PLASIM-GENIE validation against multi-model ensembles of Representative Concentration Pathways

	RCP2.6		RCP8.5	
	CMIP5	PLASIM-GENIE	CMIP5	PLASIM-GENIE
Warming (°C)	1.0 ± 0.4 (0.3, 1.7)	0.7 ± 0.2 (0.4, 1.0)	3.7 ± 0.7 (2.6, 4.8)	3.6 ± 0.6 (2.6, 4.4)
CO_2 (ppm)		402 ± 19 (373, 429)	985 ± 97 (794, 1142)	1010 ± 110 (829, 1185)
AMOC (% change)		−6 ± 10 (−17, 4)	(−60, −15)	−32 ± 12 (−54, −16)
Surface pH (pH)	−0.07 ± 0.001	−0.04 ± 0.01 (−0.069, −0.028)	−0.33 ± 0.003	−0.33 ± 0.04 (−0.41, −0.27)
Productivity (%)	−2.0 ± 4.1	−2.7 ± 1.2 (−4.8, −1.2)	−8.6 ± 7.9	−15.1 ± 4.1 (−21.7, −7.43)

Data are expressed as 2090–1990 decadal anomalies except for CO_2, which is the 2100 concentration, and PLASIM-GENIE productivity, which is the 2105–2005 anomaly. The 1990 PLASIM-GENIE baselines are 30-year averages (1976–2005) except for ocean pH and productivity (where annual averages are used for all analysis). Ensembles are summarized as mean ± 1 s.d. (5th and 95th percentiles), except for CMIP5 CO_2 and AMOC, where the bracketed ranges represent 11-member and 10-member ensemble spreads respectively.

Table 2 | PLASIM-GENIE summary confidence intervals of the E3ME-FTT-GENIE-1 scenarios

	Current policies	2P0C policies	1P5C policies
Peak decadal warming (°C)	(2.54, 3.12, **4.18**, 5.17, 5.47)	(1.09, 1.19, **1.52**, 1.95, 2.02)	(1.04, 1.11, **1.40**, 1.74, 1.85)
Peak annual CO_2 (ppm)	(649, 703, **863**, 996, 1048)	(394, 405, **446**, 485, 493)	(381, 391, **429**, 458, 468)
Min decadal AMOC (%)	(33, 44, **68**, 80, 87)	(43, 76, **83**, 90, 95)	(51, 74, **84**, 90, 94)
Max annual surf acidification (pH)	(−0.50, −0.47, **−0.39**, −0.31, −0.27)	(−0.22, −0.19, **−0.15**, −0.12, −0.10)	(−0.19, −0.17, **−0.14**, −0.10, −0.09)
2100 decadal warming (°C)	(2.54, 3.12, **4.18**, 5.17, 5.47)	(0.73, 1.10, **1.41**, 1.81, 1.87)	(0.63, 0.97, **1.24**, 1.61, 1.67)
2105 annual CO_2 (ppm)	(649, 703, **863**, 996, 1,048)	(371, 382, **415**, 445, 453)	(357, 367, **394**, 416, 427)
2100 decadal AMOC (%)	(33, 45, **69**, 83, 91)	(43, 79, **90**, 102, 104)	(52, 82, **92**, 101, 107)
2105 annual surf acidification (pH)	(−0.50, −0.47, **−0.39**, −0.31, −0.27)	(−0.19, −0.17, **−0.13**, −0.10, −0.09)	(−0.16, −0.15, **−0.11**, −0.09, −0.08)
2105 annual productivity (%)	(−33.7, −24.3, **−13.8**, −4.6, −3.5)	(−9.5, −5.0, **−3.0**, −1.1, −0.8)	(−5.7, −4.1, **−2.2**, −0.7, −0.1)
Bias corrected peak warming (°C)		(1.39, 1.49, **1.82**, 2.25, 2.32)	(1.34, 1.41, **1.70**, 2.04, 2.15)
Bias corrected 2100 warming (°C)		((1.03, 1.40, **1.71**, 2.11, 2.17)	(0.93, 1.27, **1.54**, 1.91, 1.97)

Minima, 5th percentile, median (bold), 95th percentile and maxima of the 69-member ensembles. Warming, AMOC and acidification are expressed relative to a 30-year average baseline centred on 1870. Productivity is the 2105–2005 anomaly. The 0.3 °C bias correction under strong mitigations is implied by the RCP2.6 CMIP5 comparison (Table 1).

to CMIP5, underestimating the warmest ensemble members (Table 1). We therefore apply 0.3 °C to bias-correct warming estimates in the rapid decarbonization scenarios (Table 2).

Our future simulations are forced with emissions from policy scenarios of the simulation-based integrated assessment model E3ME-FTT-GENIE[18]. The E3ME macroeconomic model differs fundamentally from the equilibrium models more usually used to assess climate policy by representing realistic (non-optimal) behaviour based on empirical relationships, and by relaxing the constraint of a fixed money supply. Investment in renewables therefore can in principle generate economic stimulus, for instance, through increased employment.[19] Furthermore, the framework is suited to flexible application of a range of policy implementations that are not limited to a carbon tax, including regulations, subsidies, focussed taxation policies and public procurement. The model contains a bottom-up representation of technological diffusion in multiple sectors (FTT) and is connected to a climate carbon cycle model (GENIE) with a single-layer atmosphere. We consider three scenarios: (1) CP, current policy[18,20] (2) 2P0C[18,20], which is a rapid decarbonization policy aimed at avoiding 2 °C peak warming with 75% confidence (according to GENIE) and (3) 1P5C (see Methods), which represents our most optimistic set of policy assumptions, avoiding 1.5 °C peak warming with 50% confidence.

Time series for the PLASIM-GENIE ensembles forced with the three policy scenarios are illustrated in Fig. 1, and ensemble distributions are summarized in Table 2. We note that the time series of ensemble median values do not correspond to fixed simulations, so the distribution of peak decadal warming (Table 2) shows slightly higher values as individual trajectories cross owing to decadal variability. Steady-state decadal variability of mean surface temperature in PLASIM-GENIE is ±0.08 °C (one standard deviation).

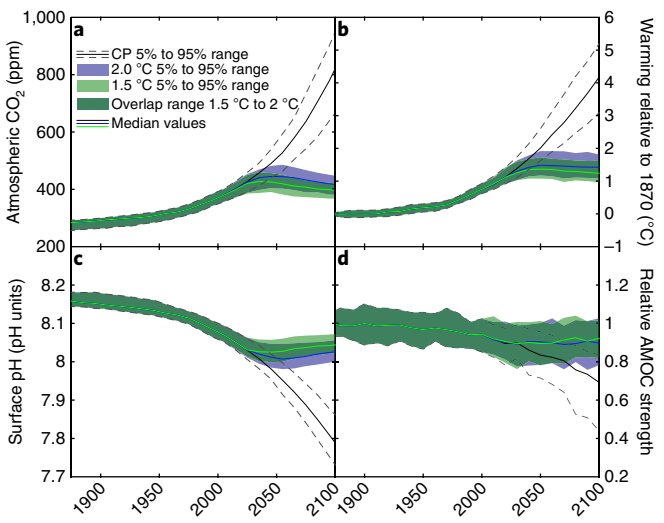

Fig. 1 | Summary time series of the 69-member CP, 2P0C and 1P5C E3ME-FTT-GENIE emissions-forced PLASIM-GENIE ensembles. a, Atmospheric CO_2. **b**, Warming relative to 1870. **c**, Surface pH. **d**, Relative AMOC strength.

Small differences in assumptions can make significant differences to cumulative emissions budgets under strong mitigation, noting

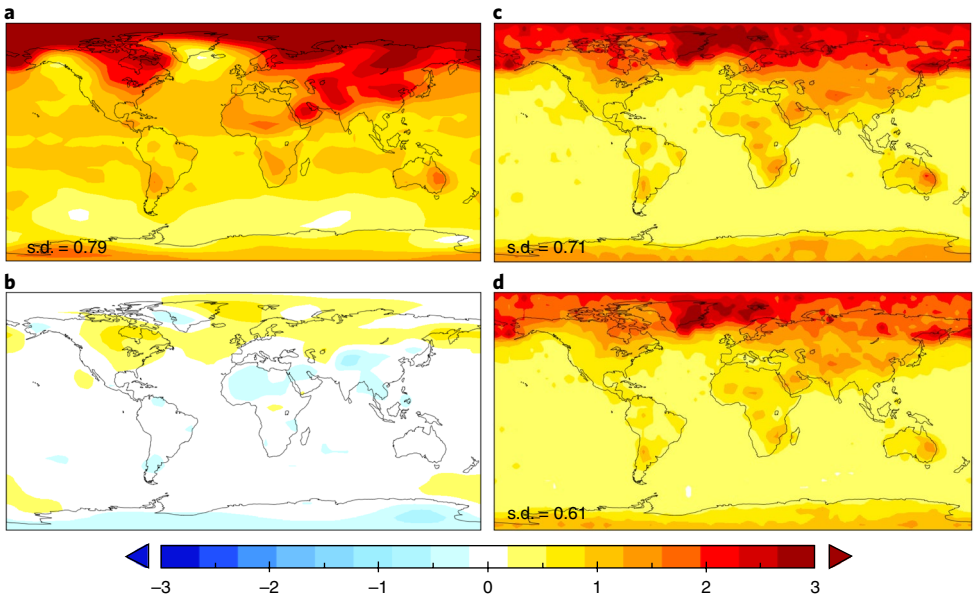

Fig. 2 | DJF surface air temperature scaling patterns and uncertainty. a, Scaling pattern from the 1P5C scenario, the ensemble mean anomaly (2086–2095 minus 1976–2005, in °C) normalized per 1 °C warming. **b**, The pattern uncertainty, the difference in patterns derived from 1P5C and CP ensembles (1P5C normalized mean - CP normalized mean). **c**,**d**, Ensemble variability is derived for both 10-year (2086–2095) (**c**) and 30-year (2076–2105) (**d**) patternsto help isolate the contributions of decadal variability and parametric uncertainty. Ensemble variability is calculated by normalizing eachensemble member per 1 °C warming and calculating the root mean square (RMS) difference with respect to the mean pattern (**a**).

that 0.1 °C incremental warming is equivalent to approximately 50 gigatonnes of carbon.[4] Here, we consider both maximum and end-of-century change, as the former is most relevant for impact assessment and most consistent with the text of the Paris Agreement, with change expressed relative to a pre-industrial (1856–1885) baseline taken from ensembles of 1805–2105 AD transient simulations. RCP2.6 non-CO_2 forcing is applied for both mitigation scenarios, and RCP8.5 non-CO_2 forcing for the CP scenario.

Bias-corrected median peak warming estimates (Table 2) are 1.82 °C (2P0C) and 1.70 °C (1P5C), and 2100 estimates are 1.71 °C and 1.54 °C. Correlations suggest an increasing relative contribution of carbon-cycle processes to warming under rapid decarbonization (Supplementary Table 1). The response of the maximum value of Atlantic Meridional Overturning Circulation (AMOC) in the mitigation scenarios is notable. The simulated expected peak weakening to 84% of pre-industrial values (Table 2) arises from natural variability (steady-state decadal variability is 0.9 Sv); the median response through the century is steady (Fig. 1). However, in one 1P5C and two 2P0C simulations the AMOC reduces to about 50% of its present-day strength. We therefore cannot rule out significant AMOC weakening under mitigation, but note the suggestion of a reduction in the probability of this unlikely event under accelerated decarbonization.

We now consider the mean climate change patterns for a range of impact-relevant climate stressors: decadal December, January and February (DJF) surface air temperature (Fig. 2a), decadal June, July and August (JJA) precipitation (Fig. 3a), annual surface ocean acidity (Fig. 4a) and annual marine primary productivity (Fig. 4d). Patterns are 1P5C ensemble averages of (2090 minus 1990) change, expressed per 1 °C mean ensemble warming. The mean patterns of changes of temperature and precipitation are broadly consistent with CMIP5 projections. Changes in pH (Fig. 4a) result from increased concentrations of dissolved CO_2 and the associated reduction in carbonate ion concentrations is approximately uniform across the surface ocean, except in the Arctic where amplified CO_2 uptake is apparent under melting sea ice[21]. This pattern is robust, explaining more than 95% of the variability in the ensemble (quantified through singular vector decomposition); a similar robust pattern of acidification was found in CMIP5[17]. Changes in primary productivity (Fig. 4d) are dominated by large reductions of up to about 10% per °C of warming that are simulated in the equatorial Pacific Ocean. Large reductions are also simulated in the mid-latitude Pacific and Indian oceans, and in the equatorial and high-latitude Atlantic Ocean. Despite the simplified ecosystem model[22], the patterns and magnitudes of productivity change are consistent with CMIP5 analysis; in RCP8.5, decreases of up to 30%–50% are simulated in these regions[17], attributed to increased ocean stratification and slowed circulation, with consequent reductions in nutrient supply[23]. Increases in productivity are apparent in the Arctic and in parts of the Southern and Indian oceans, where the increases are probably attributable to increased nutrient supply[24]. In stark contrast to pH, the pattern of productivity change explains only 20% of ensemble variability.

The ensemble projections are now used to quantify spatio-temporal uncertainty by evaluating the adequacy of the approximations made in "pattern scaling"[25], a widely used approach to estimating climate fields for impacts evaluation. In pattern scaling, an average climate response is calculated, typically as a multi-decadal average pattern of change. The pattern, normalized per °C global mean warming, is then scaled as appropriate for scenarios of interest. The strengths and limitations of pattern scaling, including modified approaches, have recently been reviewed[26]. It is known to be less accurate under strong mitigation[27].

Figures 2b, 3b, 4b and 4e plot the normalized mean field difference (1P5C minus CP), capturing non-linear scenario-dependent feedbacks, and examining the pattern-scaling approximation of a scenario-invariant pattern. The temperature pattern differences reveal modest changes, for instance in the northern Atlantic, where the stronger AMOC leads to relatively warmer temperatures under mitigation. The largest precipitation pattern differences are associated with the Indian and southeast Asian monsoons. The magnitudes of pH change patterns are very different in the two scenarios, approximately −0.1 pH unit per °C under current policy and −0.03 per °C for rapid decarbonization. This difference reflects the different response times of pH and temperature to changing CO_2. The 2090 temperature is influenced by cumulative excess CO_2 but the surface pH in 2090 is

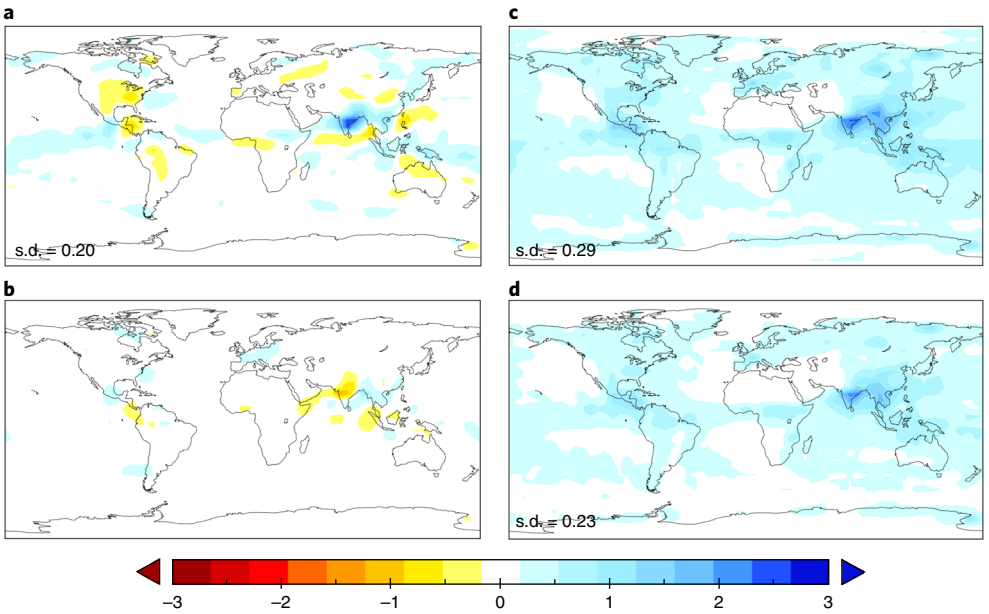

Fig. 3 | JJA precipitation scaling patterns and uncertainty. a, Scaling pattern from the 1P5C scenario, the ensemble mean anomaly (2086–2095 minus 1976–2005, in mm per day) normalized per 1 °C warming. **b**, The pattern uncertainty, the difference in patterns derived from 1P5C and CP ensembles (1P5C normalized mean -CP normalized mean). **c**,**d**, Ensemble variability is derived for both 10-year (2086–2095) (**c**) and 30-year (2076–2105) (**d**) patterns to help isolate the contributions of decadal variability and parametric uncertainty. Ensemble variability is calculated by normalizing each ensemble member per 1 °C warming and calculating the root mean square (RMS) difference with respect to the meanpattern (**a**).

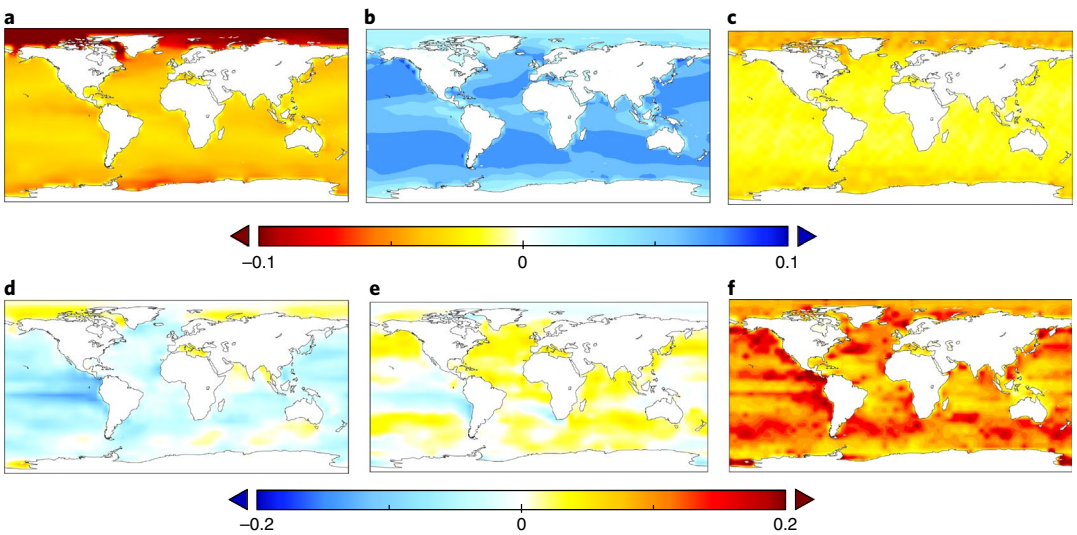

Fig. 4 | Ocean stressor scaling patterns and uncertainty. a, Surface pH scaling pattern, 1P5C ensemble mean anomaly (2105-2005), pH units per 1 °C warming. **b**, 1P5C minus CP pH scaling pattern difference. **c**, Ensemble variability of pH, calculated by normalizing each ensemble member per 1 °C warming and calculating the RMS difference with respect to the mean pattern. **d**, Marine productivity scaling pattern, 1P5C ensemble meananomaly (2105–2005), fractional change per 1 °C warming. **e**, 1P5C minus CP productivity scaling pattern difference. **f**, Ensemble variability of productivity, calculated by normalizing each ensemble member per 1 °C warming and calculating the RMS difference with respect to the mean pattern. All data are annually averaged.

determined by 2090 CO_2 with no significant lag; mitigation acts at the timescale of natural CO_2 sinks to reduce acidification impacts on the surface ocean. In contrast, the patterns of change of marine productivity in the two scenarios are spatially different, with amplified relative reductions in the Atlantic, Indian and Southern oceans, and a reduced relative reduction in the equatorial Pacific Ocean.

The most important error when using pattern scaling arises from the neglect of variability. This emerges from two distinct sources, the neglect of model uncertainty and the neglect of natural variability, both of which alter the pattern of change itself. It is well established that natural variability, which has a magnitude that differs with location, is a critical limiting factor for the accuracy of climate projections and impact evaluation[28]. If we assume that the spread of climate model outputs encompasses all possible realities, then model error can be estimated by applying the patterns from different climate models to test the robustness of the impacts that result. However, internal variability is generally not considered, and pattern scaling impacts are derived from climate means. Under strong mitigation we argue that this neglect may be inappropriate. The signal-to-noise ratio in strong mitigation scenarios is of the order of one and, for instance, decadal variability will be a significant contributor to the uncertainty in determining peak (around 2050 AD) climate change.

In the final columns of Figs. 2, 3 and 4, each 1P5C simulation anomaly field is normalized by its respective warming, and the RMS ensemble variability about the 1P5C scenario mean is plotted. For the climate fields (Figs. 2 and 3), a comparison of the variability about the mean field 30-year averages (predominantly parametric uncertainty) and 10-year averages (internal and parametric uncertainty) relative to a 30-year baseline indicates that the two sources of variability are comparable in amplitude. For the ocean impact fields (Fig. 4) the variability is derived from annual averages. In all fields, the uncertainties in the patterns (1P5C minus CP) are dominated by the variability about the pattern (right panels). The uncertainties often dominate even the mean response. For instance, in parts of the Arctic, RMS uncertainty of about 3 °C per °C of warming compares to a mean signal of around 3 °C (Fig. 2, Supplementary Table 2), while RMS uncertainty of precipitation is comparable to the mean signal in monsoon regions (Fig. 3, Supplementary Table 2). Simulations forced by current-policy emissions are associated with significantly lower fractional uncertainty (Supplementary Table 2), reflecting an increased signal-to-noise ratio, and demonstrating that the assumptions of pattern scaling are well justified under high-emission scenarios.

The implications of our findings for policy-making are important: if policy and market-based responses to climate change are sufficient to uphold the level of ambition of the Paris Agreement, climate change impacts could still be of large amplitude in sensitive regions such as the Arctic. However, in these scenarios, uncertainties from model error and internal variability can dominate expected mean patterns. Consequently, we argue that a paradigm shift in impact evaluation is now essential to support decision-making. Estimates based on mean patterns of change will be insufficient. Instead, statistical methodologies developed to address non-linear spatio-temporal feedbacks[29] will need to be extended to high-complexity models. Limiting the increase in (multi-decadal) global average temperature above pre-industrial values to 1.5 °C appears still to be possible, but results in a world where the superposition of climate change onto natural variability is key to understanding impacts on ecosystems, biodiversity, ice sheets and permafrost stability among others.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41558-018-0197-7.

Received: 31 October 2017; Accepted: 15 May 2018;
Published online: 25 June 2018

References

1. *Adoption of the Paris Agreement* FCCC/CP/2015/L.9/Rev.1 (UNFCCC, 2015);http://unfccc.int/resource/docs/2015/cop21/eng/l09r01.pdf
2. Anderson, K. & Peters, G. The trouble with negative emissions. *Science* **354**, 182–183 (2016).
3. Friedlingstein, P. et al. Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *J. Clim.* **27**, 511–526 (2014).
4. Allen, M. R. et al. Warming caused by cumulative carbon emissions towards the trillionth tonne. *Nature* **458**, 1163–1166 (2009).
5. Ehlert, D. & Zickfeld, K. What determines the warming commitment after cessation of CO_2 emissions? *Environ. Res. Lett.* **12**, 015002 (2017).
6. Steinacher, M., Joos, F. & Stocker, T. F. Allowable carbon emissions lowered by multiple climate targets. *Nature* **499**, 197–201 (2013).
7. Millar, R. J. et al. Emission budgets and pathways consistent with limiting warming to 1.5 °C. *Nat. Geosci.* **10**, 741–747 (2017).
8. Rogelj, J. et al. Scenarios towards limiting global mean temperature increase below 1.5° C. *Nat. Clim. Change* **8**, 325–332 (2018).
9. Holden, P. B. et al. PLASIM–GENIEv1.0: a new intermediate complexity AOGCM. *Geosci. Mod. Dev.* **9**, 3347–3361 (2016).
10. Geoffroy, O. et al. Transient climate response in a two-layer energy-balance model. Part I: Analytical solution and parameter calibration using CMIP5 AOGCM experiments. *J. Clim.* **26**, 1841–1857 (2013).
11. Gregory, J. M., Andrews, T. & Good, P. The inconstancy of the transient climate response parameter under increasing CO_2. *Phil. Trans. R. Soc. A* **373**, 20140417 (2015).
12. Winton, M., Takahashi, K. & Held, I. M. Importance of ocean heat uptake efficacy to transient climate change. *J. Clim.* **23**, 2333–2344 (2010).
13. Williamson, D. et al. History matching for exploring and reducing climate model parameter space using observations and a large perturbed physics ensemble. *Clim. Dynam.* **41**, 1703–1729 (2013).
14. IPCC *Climate Change 2013: The Physical Science Basis* (eds Stocker, T. F. et al.) (Cambridge Univ.Press, 2013).
15. Meinshausen, M. et al. The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change* **109**, 213–241 (2011).
16. Cheng, W., Chiang, J. C. H. & Zhang, D. Atlantic Meridional Overturning Circulation (AMOC) in CMIP5 models: RCP and historical simulations. *J. Clim.* **26**, 7187–7197 (2013).
17. Bopp, L. et al. Multiple stressors of ocean ecosystems in the 21st century: projections with CMIP5 models. *Biogeosciences* **10**, 6225–6245 (2013).
18. Mercure, J.-F. et al. Environmental impact assessment for climate change policy with the simulation-based integrated assessment model E3ME-FTT-GENIE. *Energy Strategy Rev.* **20**, 195–208 (2018).
19. Pollitt, H. & Mercure, J.-F. The role of money and the financial sector in energy-economy models used for assessing climate and energy policy. *Clim. Policy* **18**, 184–197 (2017).
20. Mercure, J.-F. et al. Macroeconomic impact of stranded fossil-fuel assets. *Nat. Clim. Change* https://doi.org/10.1038/s41558-018-0182-1 (2018).
21. Yamamoto, A., Kawamiya, M., Ishida, A., Yamanaka, Y. & Watanabe, S. Impact of rapid sea-ice reduction in the Arctic Ocean on the rate of ocean acidification. *Biogeosciences* **9**, 2365–2375 (2012).
22. Ridgwell, A. et al. Marine geochemical data assimilation in an efficient Earth System Model of global biogeochemical cycling. *Biogeosciences* **4**, 87–104 (2007).
23. Steinacher, M. et al. Projected 21st century decrease in marine productivity: a multi-model analysis. *Biogeosciences* **7**, 979–1005 (2010).
24. Rykaczewski, R. R. & Dunne, J. P. Enhanced nutrient supply to the California Current Ecosystem with global warming and in creased stratification in an Earth system model. *Geophys. Res. Lett.* **37**, L21606 (2010) .
25. Santner, B. D., Wigley, T. M. L., Schlesinger, M. E. & Mitchell, J. F. B. *Developing Climate Scenarios from Equilibrium GCM Results* (Max-Planck-Institut fuer Meteorologie, 1990).
26. Tebaldi, C. & Arblaster, J. M. Pattern scaling: its strengths and limitations, and an update on the latest model simulations. *Climatic Change* **122**, 459–471 (2014).
27. Wu, P., Wood, R., Ridley, J. & Lowe, J. Temporary acceleration of the hydrological cycle in response to a CO_2 rampdown. *Geophys. Res. Lett.* **37**, L12705 (2010).
28. Deser, C., Knutti, R., Solomon, S. & Phillips, A. S. Communication of the role of natural variability in future North American climate. *Nat. Clim. Change* **2**, 775–779 (2012).
29. Holden, P. B., Edwards, N. R., Garthwaite, P. H. & Wilkinson, R. D. Emulation and interpretation of high-dimensional climate outputs. *J. Appl. Stat.* **42**, 2038–2055 (2015).

Acknowledgements

We acknowledge C-EERNG and Cambridge Econometrics for support, and funding from EPSRC (to J.-F.M., fellowship number EP/ K007254/1); the Newton Fund (to J.-F.M., P.S. and J.E.V., EPSRC grant number EP/N002504/1 and ESRC grant number ES/N013174/1), NERC (to N.R.E., P.H. and H.P., grant number NE/P015093/1), CONICYT (to P.S.), the Philomathia Foundation (to J.E.V.) and Horizon 2020 (to H.P. and J.-F.M., the Sim4Nexus project).

Author contributions

P.B.H., N.R.E. and R.D.W. designed and coordinated the Earth system modelling. H.P., J.-F.M. and N.R.E. designed and coordinated the energy-economy modelling. P.B.H., N.R.E., R.D.W. and H.P. wrote the article with contributions from all. P.B.H. performed the PLASIM-GENIE simulations. U.C. performed the E3ME-FTT simulations. All authors developed model components and/or provided scientific support: P.B.H. (ESM coupling), K.F. and F.L. (atmosphere), N.R.E. (ocean), A.R. (biogeochemistry), H.P. and J.F.M. (energy-economic), P.S. and J.-F.M. (power sector), A.L. and J.-F.M. (transport sector), F.K. and J.-F.M. (household heating), J.E.V. (geopolitics) and R.D.W. (statistics).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41558-018-0197-7.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to P.B.H.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

PLASIM-GENIE is a coupling of the intermediate-complexity spectral atmosphere model PLASIM[30] to the Grid-Enabled Integrated Earth system model GENIE[31]. The coupling and climatology are described in detail in ref. [9]. PLASIM-GENIE is not flux-corrected; the moisture flux correction required in the original tuning[9] was removed during the history-matching calibration (see below). We here apply PLASIM-GENIE with carbon-coupled biosphere modules BIOGEM and ENTS, described in ref. [31] for the energy–moisture balance atmosphere configuration. We apply BIOGEM with the default Michaelis–Menton phosphate-limited productivity scheme[22]. The carbon-cycle model has been extensively validated through model intercomparisons[32,33].

Important neglects of the PLASIM-GENIE carbon cycle are anthropogenic land-use change, peat and permafrost. These omissions tend to overstate the terrestrial carbon sink (by overstating natural forest) and they neglect potentially significant terrestrial sources (from peat and permafrost). We note that the history-matching calibration is designed to subsume such structural deficiencies (here, for instance, into CO_2 fertilization and soil respiration).

PLASIM-GENIE is freely available. Please contact the corresponding author for information.

Atmosphere–ocean gearing. PLASIM-GENIE simulates approximately 2.5 years per central processing unit (CPU) hour, so that 2,000-year spin-ups take one month of computing. To enable the exploration of parameter space, the implementation of an atmosphere–ocean gearing approach was required. The spin-up simulation time is determined by the ocean timescale, but the simulation speed of the model is determined by the atmosphere, which uses approximately 99% of the CPU demands of the physical model. In gearing mode, applied only to equilibrium spin-ups, the model alternates between a conventionally coupled mode (for 1 year) and a fixed-atmosphere mode (for 9 years), reducing spin-up CPU time by an order of magnitude. During the conventional coupling mode, atmosphere–ocean coupling variables are accumulated and saved as daily averages. These variables comprise energy fluxes, moisture fluxes and wind stresses. During the fixed-atmosphere phase, the atmospheric variables are kept constant and these daily averaged fluxes are applied to the ocean. Latent heat, sensible heat and longwave radiation ocean heat loss are recalculated at every atmosphere time step during the fixed atmosphere phase, when energy conservation is therefore not imposed. This is necessary for numerical convergence because these fluxes depend on ocean temperature, which evolves during the fixed atmosphere phase. Evaporation is not recalculated during the fixed atmosphere phase to ensure moisture conservation. Atmosphere–ocean-geared spin-up states are consistent with the standard model, as demonstrated by smooth spun-on historical transient simulations in all ensemble members, though we note that rapid (sub-decadal) and modest (a few Sv) AMOC adjustments are seen in some simulations, arising from different interannual variability.

Experimental design. Each model configuration was spun-up with a 2,000-year atmosphere–ocean-geared quasi-equilibrium pre-industrial simulation, with atmospheric CO_2 relaxed to 278 parts per million (ppm). Simulations were continued as emissions-forced historical transient simulations (atmosphere–ocean gearing off, CO_2 freely evolving). Historical forcing (1805 to 2005) comprised anthropogenic CO_2 emissions and non-CO_2 radiative forcing. Fossil fuel, cement and gas-flaring emissions were prescribed from CMIP5 (https://cmip.llnl.gov/cmip5/forcing.html) and were combined with Integrated Science Assessment Model land-use change emissions[34] from the HYDE land-use dataset[35]. Non-CO_2 forcing data was taken from ref. [15] implemented in PLASIM-GENIE as effective CO_2. Future (2005–2105) emissions were taken from the E3ME-FTT-GENIE scenarios, scaled by 9.82/8.62, to match estimated 2015 total emissions[36], accounting for sources not represented in E3ME. Future land-use change emissions and non-CO_2 radiative forcing were taken from RCP2.6 (the 1P5C and 2P0C scenarios) and RCP8.5 (the CP scenario).

History-matched ensemble. Carefully designed ensembles of simulations are central to our approach to quantifying Earth system uncertainties. We applied a 'history matching' calibration strategy[13,37], sampling throughout high-dimensional model input space to identify model configurations that are capable of producing reasonable simulations in the PLASIM-GENIE Earth system model, and then running the plausible configurations forward to characterize uncertainty about the future. Each configuration is required only to provide a 'plausible' simulation[38], thereby avoiding the introduction of bias through over-fitting[39]. A configuration is ruled out only if it is inconsistent with an observation, allowing for the imperfections of both model and data. Thus, the history-matching philosophy generates simulations that encompass the full range of realistic dynamical feedbacks implemented in the model[40].

In PLASIM-GENIE, identifying large numbers of history-matched configurations would be prohibitively demanding computationally. We render the problem tractable by using emulators[41] to search throughout model input space. The emulators are trained on a sequence of preliminary ensembles amounting to 1.9 million years of climate simulation in total (940 completed simulations). The process produced 69 model variants, each validated by simulation, having considered hundreds of millions of randomly sampled parameter configurations in the emulator. The final models all adequately simulate ten key global-scale observational targets including surface air temperature, vegetation and soil carbon, Atlantic, Pacific and Southern ocean circulation measures, dissolved O_2 and calcium carbonate flux, and transient temperature and CO_2 changes (Supplementary Table 4).

For the purposes of the history matching, the simulator (here applied to the pre-industrial spin-up state) can be considered as a function that maps from 32 input parameters (Supplementary Table 3) to the eight different outputs (Supplementary Table 4). Our aim is to infer the input values that lead to outputs within the plausible climate ranges as defined in Supplementary Table 4. It is not possible to naively explore the simulator output over the full input parameter ranges by repeatedly evaluating the simulator, as for example, just doing one evaluation in each corner of the input space would require $2^{32} \approx 10^9$ model evaluations. Instead, we build emulators[41,42] that mimic the simulator response surface, and allow us to predict its value for any input. An initial large exploratory analysis was performed, motivated by the iterated waves approach[39]. Starting from a 100-member maxi-min Latin hypercube ensemble, sequential series of 100-member ensembles were performed, probing regions of likely plausible space by using stepwise-selected linear regression models that were continually refitted as simulations completed. This produced 940 completed simulations that we used to train the final history match. Part of the motivation for the exploratory ensemble was to develop a general understanding of the range of model responses. Most notably, it enabled us to identify regions of parameter space that satisfied the plausibility constraints without flux correction so that the associated parameter (Supplementary Table 3) could be fixed at zero for the final history match.

For the final history match, a variety of emulation approaches were considered, including stepwise regression, the LASSO[43], which is a regularized version of linear regression, and Gaussian process regression with a combination of different mean and covariance functions[44]. To determine the optimal approach for each of the eight outputs, we split the data into test and training datasets and evaluated the emulators' predictive performance (RMS error, statistical coverage), repeating the process ten times to get an average performance. The optimized emulators were used to find input values that are expected to give plausible simulations (that is, within tabulated ranges for all emulator-filtered metrics; Supplementary Table 4), to generate a sample of design points which encapsulate the uncertainty about future climate. We used an approximate Bayesian computation type approach[45], using rejection sampling to sample parameters from the prior distribution and evaluating the probability of these values leading to plausible outputs, to generate a large number of plausible future climates, considering hundreds of millions of emulator evaluations. A final 200-member candidate ensemble for the future transient simulations was then chosen using a 'greedy' design, adding points to maximize a criterion that combined the probability that the simulation would be plausible (according to the emulator), and the distance of candidate points to the other points already in the design, so as to ensure the design points fully span the 32-dimensional plausible input space.

The 200 history-matched parameter sets were applied to PLASIM-GENIE, and 183 were accepted as giving plausible pre-industrial climates in the simulator. These were spun on through the industrial period (1805 to 2005) with emissions and non-CO_2 radiative forcing. Sixty-nine simulations were selected as also having plausible climate sensitivity (2005–1870 warming between 0.6 K and 1.0 K) and carbon cycle (2005 CO_2 in the range 355 ppm to 403 ppm). These 69 model configurations were applied in the future transient ensembles.

In total, 1,140 spin-up simulations (2,000 years each) were performed with the geared model and 345 transient simulations (300 years each) with the standard model, representing approximately 15 CPU years of computing, corresponding to the CPU time needed to simulate a few decades with a CMIP5 type Earth system model.

Decarbonization policies to meet 1.5 °C and 2 °C. The E3ME-FTT-GENIE modelling framework and the particular policy scenarios used here have been described in detail elsewhere[18,20], and below we -give a summary of the policy choices taken as inputs to the modelling framework in deriving the emissions scenarios used here as input to PLASIM-GENIE. Three scenarios are used: a current policy baseline, a scenario in which there is an 75% chance of limiting peak warming to 2 °C and a scenario in which there is a 50% chance of limiting peak warming to 1.5 °C.

The model baseline is consistent with the IEA's current-policy scenario[46]. The baseline can broadly be considered as a continuation of current policy trends; existing policy remains in place and has some lagged effects that continue into the projection period, but there is no additional policy stimulus. Most policy instruments in the baseline are implicitly accounted for through the data itself (such as diffusion trends).

The 1.5 °C and 2 °C scenarios are designed as sets of policies that are added to the baseline case. In almost all countries, these policies encapsulate the measures put forward in the intended nationally determined contributions that were submitted to the Paris Conference of the Parties and complement them with other measures, to scale up the level of ambition of decarbonization. The scenarios are designed from a bottom-up perspective. Essentially, policies are added across the

full range of economic sectors sequentially until the targets are met. The 1.5 °C scenario includes all the measures in the 2 °C scenario, plus additional ones, as described below.

Many of the policies are specific to particular sectors, but two economy-wide policies are implemented:

- The first measure is an economy-wide programme of energy efficiency. Our 2 °C scenario assumes that the programmes are in line with the IEA's analysis[47] for a 450 ppm scenario (excluding houses, which are treated separately, see below). They are further scaled up by 25% for the 1.5 °C scenario.
- The second measure is a carbon tax that is applied equally across the world. The carbon tax rates rise to US$310.4 per tCO_2 and US$96.2 per tCO_2 by 2030 in the 1.5 °C and 2 °C scenarios respectively, and US$886.3 per tCO_2 and US$274.8 per tCO_2 by 2050 (at 2015 prices). The carbon taxes are applied to all industrial sectors, but not to road transport or households, where separate rates are levied (since these sectors are likely to, or already have, their own specific carbon or energy tax rates).

Building on ref. [48], the following power sector policies were added to both scenarios:

- Feed-in tariffs: 100% of the difference between the levelized cost for wind and solar power and a fixed value of US$80 per MWh (at 2015 prices) is paid by the grid to promote renewable uptake.
- Direct renewables subsidies: in most cases 50% to 60%, to provide an incentive to increase uptake, across a range of technologies (this is in addition to feed-in tariffs). The subsidies gradually decrease over time and are phased out by 2050.
- In several countries there are immediate mandates to prevent the construction of new coal capacity.

In addition, it is assumed that electricity storage technologies advance up to 2050 such that the requirement for back-up flexible generation capacity (such as oil and gas peaking plants) is limited. Combinations of policies are used to incentivize the adoption of vehicles with lower emissions[49] in both scenarios. The list includes:

- Fuel-efficiency regulations of new liquid fuel vehicles.
- A phase-out of older models with lower efficiency.
- Kick-start programmes for electric vehicles where they are not available (by public authorities or private institutions, such as municipality vehicles and taxis).
- A tax of US$50 increasing to US$150 per gCO_2 per kilometre (2015 prices), to incentivize vehicle choice.
- A fuel tax (increasing from US$0.10 in 2018 to US$1.00 per litre of fuel in 2050, at 2015 prices) to curb the total amount of driving.
- Increasing or introducing biofuel mandates between current values to between 10% and 30% (40% in Brazil) in 2050, different for every country, extrapolating IEA projections[50] for the 2 °C scenario, and to 97% in the 1.5 °C scenario.

Aviation is assumed to switch to biofuels gradually over the period 2020–2050 (faster in the 1.5 °C scenario), but total bioenergy consumption remains within 150 EJ per year.

The following policies were applied to homes in both scenarios:

- Taxes on the residential use of fossil fuels, applied globally, starting at an equivalent of US$110 per CO_2 (2015 values) and increasing by 10% per year afterwards.
- Direct capital subsidies on renewable heating systems, applied globally: minus 50% on the purchase and installation of heat pumps, solar thermal systems and modern biomass boilers, phased out between 2030 and 2050.
- Kick-start programmes for renewable heating systems where they are not available, for a limited time period of five years (such as installations in publicly owned housing stock).

In some industrial sectors in East and South East Asia, a further mandate was added to electrify sectors that are currently dependent on coal (only in the 1.5 °C scenario). Emissions from industrial processes are modelled as fixed in relation to real production levels from the relevant sector. In the baseline scenario, no efficiency improvements are assumed. In the 2 °C and 1.5 °C scenarios it is assumed that the production efficiency of process emissions improves by 3% a year over the projection period. Land-use change emissions are calculated in GENIE, with land-use change assumed to follow RCP2.6 in the mitigation scenarios and RCP8.5 in the current policy baseline.

Data availability. The data that support the findings of this study are available from the corresponding author on request.

References

30. Fraedrich, K. A suite of user-friendly climate models: hysteresis experiments. *Eur. Phys. J. Plus* **127**, 53 (2012).
31. Lenton, T. M. et al. Millennial timescale carbon cycle and climate change in an efficient Earth system model. *Clim. Dynam* **26**, 687–711 (2006).
32. Zickfeld, K. et al. Long-term climate change commitment and reversibility: an EMIC intercomparison. *J. Clim.* **26**, 5782–5809 (2013).
33. Joos, F. et al. Carbon dioxide and climate impulse response functions for the computation of greenhouse gas metrics: a multi-model analysis. *Atmos. Chem. Phys.* **13**, 2793–2825 (2013).
34. Jain, A. K., Meiyappan, P., Song, Y. & House, J. I. CO_2 emissions from land-use change affected more by nitrogen cycle, than by the choice of land cover data. *Glob. Change Biol.* **19**, 2893–2906 (2013).
35. Ramankutty, M. et al. Challenges to estimating carbon emissions from tropical deforestation. *Glob. Change Biol.* **13**, 51–66 (2007).
36. Jackson, R. B. et al. Reaching peak emissions. *Nat. Clim. Change* **6**, 7–10 (2016).
37. Craig, P. S., Goldstein, M., Seheult, A. H. & Smith, J. A. in *Case Studies in Bayesian Statistics* (Lecture Notes in Statistics, Springer, New York, 1997).
38. Edwards, N. R., Cameron, D. & Rougier, J. Precalibrating an intermediate complexity climate model. *Clim. Dynam.* **37**, 1469–1482 (2011).
39. Williamson, D. B., Blaker, A. T. & Sinha, B. Tuning without over-tuning: parametric uncertainty quantification for the NEMO ocean model. *Geosci. Model Dev.* **10**, 1789–1816 (2017).
40. Holden, P. B., Edwards, N. R., Oliver, K. I. C., T. Lenton, T. M. & Wilkinson, R. D. A probabilistic calibration of climate sensitivity and terrestrial carbon change in GENIE-1. *Clim. Dynam.* **35**, 785–806 (2010).
41. Sacks, J., Welch, W. J., Mitchell, T. J. & Wynn, H. P. Design and analysis of computer experiments. *Stat. Sci.* **4**, 409–23 (1989).
42. O'Hagan, A. Bayesian analysis of computer code outputs: a tutorial. *Reliab. Eng. Syst. Saf.* **91**, 1290–1300 (2006).
43. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B* **58**, 267–88 (1996).
44. Rasmussen, C. E. Gaussian processes in machine learning. In *Advanced Lectures on Machine Learning* (eds Bousquet, O., von Luxburg, U. & Rätsch, G.) 63–71 (Springer, Berlin, 2004).
45. Marin, J.-M., Pudlo, P., Robert, C. P. & Ryder, R. J. Approximate Bayesian Computational Methods. *Stat. Comput.* **22**, 1167–1180 (2012).
46. *World Energy Outlook 2015* (OECD/IEA, 2015); https://www.iea.org/publications/freepublications/publication/WEO2015.pdf
47. *World Energy Investment Outlook* (OECD/IEA, 2014); https://www.iea.org/publications/freepublications/publication/WEIO2014.pdf
48. Mercure, J. F., Pollitt, H., Bassi, A. M., Viñuales, J. E. & Edwards, N. R. Modelling complex systems of heterogeneous agents to better design sustainability transitions policy. *Glob. Environ. Change* **37**, 102–115 (2016).
49. Mercure, J. F., Lam, A., Billington, S. & Pollitt, H. Integrated assessment modelling as a positive science: private passenger road transport policies to meet a climate target well below 2 degrees C. Preprint at https://arxiv.org/abs/1702.04133 (2018).
50. *World Energy Outlook 2014* (OECD/IEA, 2014); https://www.iea.org/publications/freepublications/publication/WEO2014.pdf