

LETTER • OPEN ACCESS

Evaluating the accuracy of climate change pattern emulation for low warming targets

To cite this article: Claudia Tebaldi and Reto Knutti 2018 *Environ. Res. Lett.* **13** 055006

View the [article online](#) for updates and enhancements.

You may also like

- [Large-scale emulation of spatio-temporal variation in temperature under climate change](#)

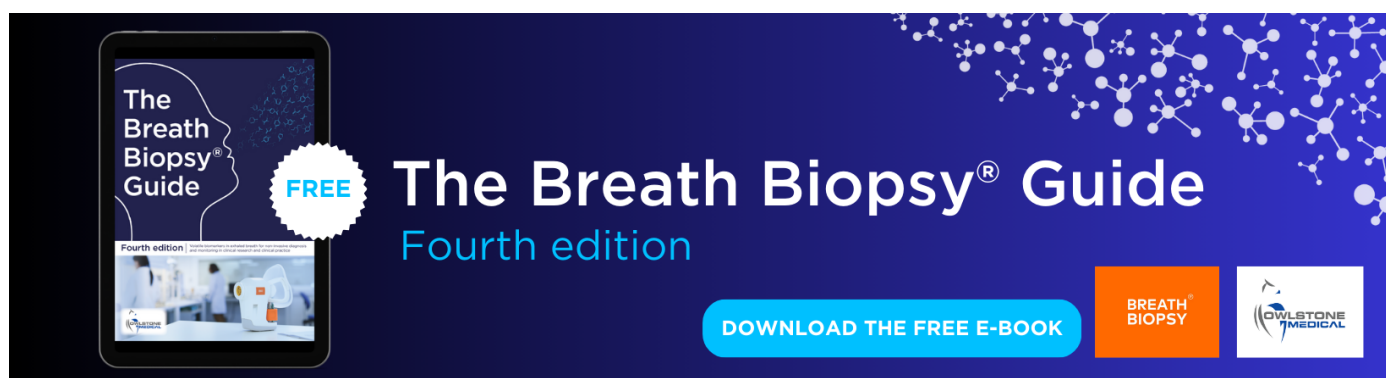
Xiao-Chen Yuan, Nan Zhang, Wei-Zheng Wang et al.

- [Heat wave exposure in India in current, 1.5 °C, and 2.0 °C worlds](#)

Vimal Mishra, Sourav Mukherjee, Rohini Kumar et al.

- [Extreme precipitation over East Asia under 1.5 °C and 2 °C global warming targets: a comparison of stabilized and overshoot projections](#)

Donghuan Li, Tianjun Zhou and Wenxia Zhang



The Breath Biopsy® Guide
Fourth edition

FREE

DOWNLOAD THE FREE E-BOOK

BREATH BIOPSY

OWLSTONE MEDICAL

Environmental Research Letters



LETTER

Evaluating the accuracy of climate change pattern emulation for low warming targets

OPEN ACCESS

RECEIVED

30 October 2017

REVISED

10 April 2018

ACCEPTED FOR PUBLICATION

18 April 2018

PUBLISHED

3 May 2018

Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Claudia Tebaldi^{1,3}  and Reto Knutti^{1,2}

¹ National Center for Atmospheric Research, Boulder, CO, United States of America

² Institute for Atmospheric and Climate Science, ETH Zurich, CH-8092 Zurich, Switzerland

³ Author to whom any correspondence should be addressed.

E-mail: tebaldi@ucar.edu

Keywords: BRACE1.5, Paris agreement, benefits of mitigation, pattern scaling, low-warming scenarios, climate model emulation

Supplementary material for this article is available [online](#)

Abstract

Global climate policy is increasingly debating the value of very low warming targets, yet not many experiments conducted with global climate models in their fully coupled versions are currently available to help inform studies of the corresponding impacts. This raises the question whether a map of warming or precipitation change in a world 1.5 °C warmer than preindustrial can be emulated from existing simulations that reach higher warming targets, or whether entirely new simulations are required. Here we show that also for this type of low warming in strong mitigation scenarios, climate change signals are quite linear as a function of global temperature. Therefore, emulation techniques amounting to linear rescaling on the basis of global temperature change ratios (like simple pattern scaling) provide a viable way forward. The errors introduced are small relative to the spread in the forced response to a given scenario that we can assess from a multi-model ensemble. They are also small relative to the noise introduced into the estimates of the forced response by internal variability within a single model, which we can assess from either control simulations or initial condition ensembles. Challenges arise when scaling inadvertently reduces the inter-model spread or suppresses the internal variability, both important sources of uncertainty for impact assessment, or when the scenarios have very different characteristics in the composition of the forcings. Taking advantage of an available suite of coupled model simulations under low-warming and intermediate scenarios, we evaluate the accuracy of these emulation techniques and show that they are unlikely to represent a substantial contribution to the total uncertainty.

Introduction

The Paris Agreement resulting from the 21st Conference Of the Parties (COP21) and the upcoming Special Report ‘Global Warming of 1.5 °C’ of the Intergovernmental Panel on Climate Change (IPCC) are stimulating research towards characterizing impacts in a world abiding by such low warming targets compared to impacts under future global warming levels of 2 °C or higher. However, very few global climate projections consistent with the low targets agreed in Paris exist (Sanderson *et al* 2017, Mitchell *et al* 2017). The impact research community could find valuable and expedient use of surrogate climate projections

approximating climate change in a 1.5 °C and 2 °C world on the basis of currently available model simulations under higher scenarios, if their accuracy was demonstrated. Established emulation techniques like simple pattern scaling (Santer *et al* 1990, Tebaldi and Arblaster 2014) have long been used to provide inputs to impact studies. An alternative approach recently documented in Herger *et al* (2015) and King *et al* (2017) consists of using transient simulations at the time when they reach a given warming target on their way to higher warming levels (from now on referred to as ‘time-shift’ approach). The accuracy of these methods, however, has seldom been quantified. More importantly, they have not been tested specifically on very

low warming/strong mitigation scenarios, and with regard to their accuracy in representing both natural variability and inter-model spread in a multi-model context.

In this study we take advantage of a set of simulations newly performed with the National Center for Atmospheric Research-Department of Energy Community Earth System Model, version 1 (CESM1) documented by Sanderson *et al* (2017). These simulations have been specifically designed to stabilize by the end of the century at a global warming of 1.5 °C and 2 °C relative to preindustrial. We use existing simulations by the same model configuration under a higher scenario from the Representative Concentration Pathway set (Moss *et al* 2010), specifically RCP4.5 (Sanderson *et al* 2018), to construct the pattern-scaling and time-shift approximations; we then test their validity in a ‘perfect model’ setup, i.e. a case where the true target is known. However, we believe that the multi-model ensemble paradigm should furnish the backdrop against which single model exercises take place. Thus, we first ask more generally how well a strong mitigation scenario, whose global temperature trajectory stabilizes during the last decades of the century at a low level of warming, can be approximated by available, more moderately mitigated trajectories. We approach this more general question within the multi-model framework provided by the simulations available from the Coupled Model Intercomparison Project phase 5 (CMIP5, Taylor *et al* 2012). From this set of simulations, we use the strongly mitigated, lowest scenario, RCP2.6 (which can be viewed as a stand-in for a low warming/high mitigation scenario of the ‘Paris kind’) and the next higher scenario RCP4.5. Therefore, we will apply the two emulation techniques to a set of models that provide simulations under both RCP4.5 and RCP2.6; we will then assess the accuracy of the emulation by comparing its error to the spread of model responses under the target scenario, RCP2.6. We also employ simulations of pre-industrial control climate (where no changes in greenhouse gases or other external forcings are imposed) to gauge the portion of the error in the estimates of the forced response to a given concentration pathway that is introduced by internal variability within each model. Similarly, when we focus on the emulation performance within CESM1, we can rely on multiple initial condition ensemble members under each scenario to assess the role of internal variability.

We focus on the two key variables of mean near-surface temperature and precipitation, averaged annually. We analyze the emulation techniques’ performance for geographic patterns of multi-decadal change, indicative of the model response to external forcings, and for the interannual variability that is superimposed on it.

Methods

As background to the approach and results presented here, we performed a thorough exploratory analysis of the value of using more than simply the closest scenario available for emulating our target. For example, we tried to include patterns from a stabilized scenario (the long-term extension of RCP4.5 producing a stationary climate over the 22nd and 23rd centuries); we also utilized results from an idealized 1%/yr increasing CO₂ scenario, providing single-forcing patterns in addition to those resulting from the mixed-forcings (CO₂, aerosols, land-use change) that characterize RCPs. We also tested the performance of rescaling RCP8.5 rather than RCP4.5. We concluded that the simple rescaling or time-shift of RCP4.5 provided the most accurate emulation of the lower scenarios, RCP2.6. Therefore, we focus the description of the data and methods on this simple approach, but the readers may be interested in knowing that those attempts were made, resulting in no significant improvement in the metrics of performance described below. Detailed results are available from the first author upon request.

We use 23 models from the CMIP5 archive providing pre-industrial control (piControl), historical, RCP2.6 and RCP4.5 simulations (one ensemble member per model, see table S1 for a list of the models and their affiliations). All models’ temperature and precipitation output is regridded to a common rectangular grid, whose gridboxes are approximately 2.5° in longitude/latitude (equivalent to about 250 km at the equator). In this CMIP5 context we targeted the emulation of RCP2.6 temperature and precipitation changes at the end of the 21st century (2081–2100) by the use of RCP4.5 simulations, model by model. We use differences across the multi-model ensemble and within each model’s piControl experiments as measures of unavoidable uncertainty in the emulated quantities (either from model structural uncertainty or internal variability, respectively, assuming the latter is not significantly different between the piControl and the future simulations). Similarly, when using CESM1 experiments, we use RCP4.5 (a 10 member initial condition ensemble) to approximate temperature and precipitation changes under either the 1.5 °C or 2.0 °C simulations, which were conducted using the same set of 10 initial condition ensemble members (i.e. the three experiments share the same 10 historical runs, up to 2005 when the different scenario forcings start being applied). For these CESM1 experiments the role of internal variability is directly quantifiable by computing the spread around the forced response (20 year average changes aggregated over all members) from the individual ensemble members. These members differ from each other only by small perturbations in their initial conditions, and therefore give a direct measure

of the variability in the outcome only due to the natural noise in the system.

The two methods of emulations whose performance we evaluate, simple pattern scaling and time-shift, have been identified for example by James *et al* (2017) as methods to approximate low-warming climate scenarios. In simple pattern scaling, a geographical snapshot of change *per degree of global mean warming* is derived from an available scenario simulation: a 20 year mean of the pattern of change at the end of the century (2081–2100 minus 1986–2005 mean change, grid-point by grid-point) is computed and is normalized by dividing it by the corresponding global average temperature change during the same period. The pattern of change under the target scenario is then emulated by multiplying the normalized pattern by the global average temperature change under the target scenario. We do so model by model for the CMIP5 case, or ensemble member by ensemble member in the CESM1 case, using the actual global temperature change produced by each model under the target scenario. Most pattern scaling applications have to approximate also the global mean warming of the target scenario, and they usually do so using a simple climate model that can be run at low computational cost, like an energy balance model (Meinshausen *et al* 2011). Here we assume that the emulated global mean has a small error compared to the approximation of the regional outcomes, and we exploit the availability of the true global warming signal from the target scenarios.

For the time shift approach applied to the CMIP5 models, we identify the times in their simulations of RCP4.5 when global average temperature reaches the same value as the end of the century (2081–2100 average) global average temperature from the corresponding simulations under RCP2.6. We then consider a 20 year window around that time as our surrogate twenty-year's worth of simulated climate under RCP2.6. Similarly, we take the 20 year window centered around the times at which the RCP4.5 simulations in the CESM1 ensemble reach the same anomaly as the 1.5 °C and 2 °C simulations do by the end of the 21st century.

To assess the accuracy of the emulation, geographic patterns of twenty-year average temperature and percent precipitation change from the emulated output are compared to the 'true' patterns computed from the target experiments. We use root mean squared errors (RMSEs), where the mean is performed by area-weighting each grid point by the cosine of its latitude, in order not to over-weight the grid-point errors at high latitudes. In order to assess the resulting values of these error metrics, we compare them to the structural model differences. We quantify the latter by computing RMSEs between each possible pair of true RCP2.6 patterns from the multi-model ensemble. Further, we want to compare the emulation error to the variations introduced by internal variability. We therefore compute RMSEs from pairs of 20 year

patterns computed along the multi-century piControl runs for the CMIP5 models (since they do not provide large ensembles of future simulations). In the case of CESM1, we simply compute RMSEs between each pair of true target patterns (either from the 1.5 °C or from the 2.0 °C experiment) among the 10 ensemble members available. The distributions of values of RMSEs from the different sources (emulation error, model differences, internal variability) are shown through their histograms plotted on a common axis, by which the size of the error in mean and range are easily compared.

Next, we consider the behavior of a 20 year series of fields of change, in order to assess how our emulations replicate interannual (year-to-year) variability. For simple pattern scaling, we take the pattern of forced change that the method delivers and we superimpose an estimate of interannual variability. Our estimate of year-to-year variability is obtained by considering the 20 year time series of fields of change from RCP4.5 after subtracting the 20 year average change, used to derive the emulation. We are therefore assuming that inter-annual variability around the pattern of forced change is similar across the scenarios. In the case of the time-shift method, we have a time series of 20 year fields of change by construction, since we are simply isolating a twenty-year window during the simulations whose global average temperature change is the same as the target scenarios.

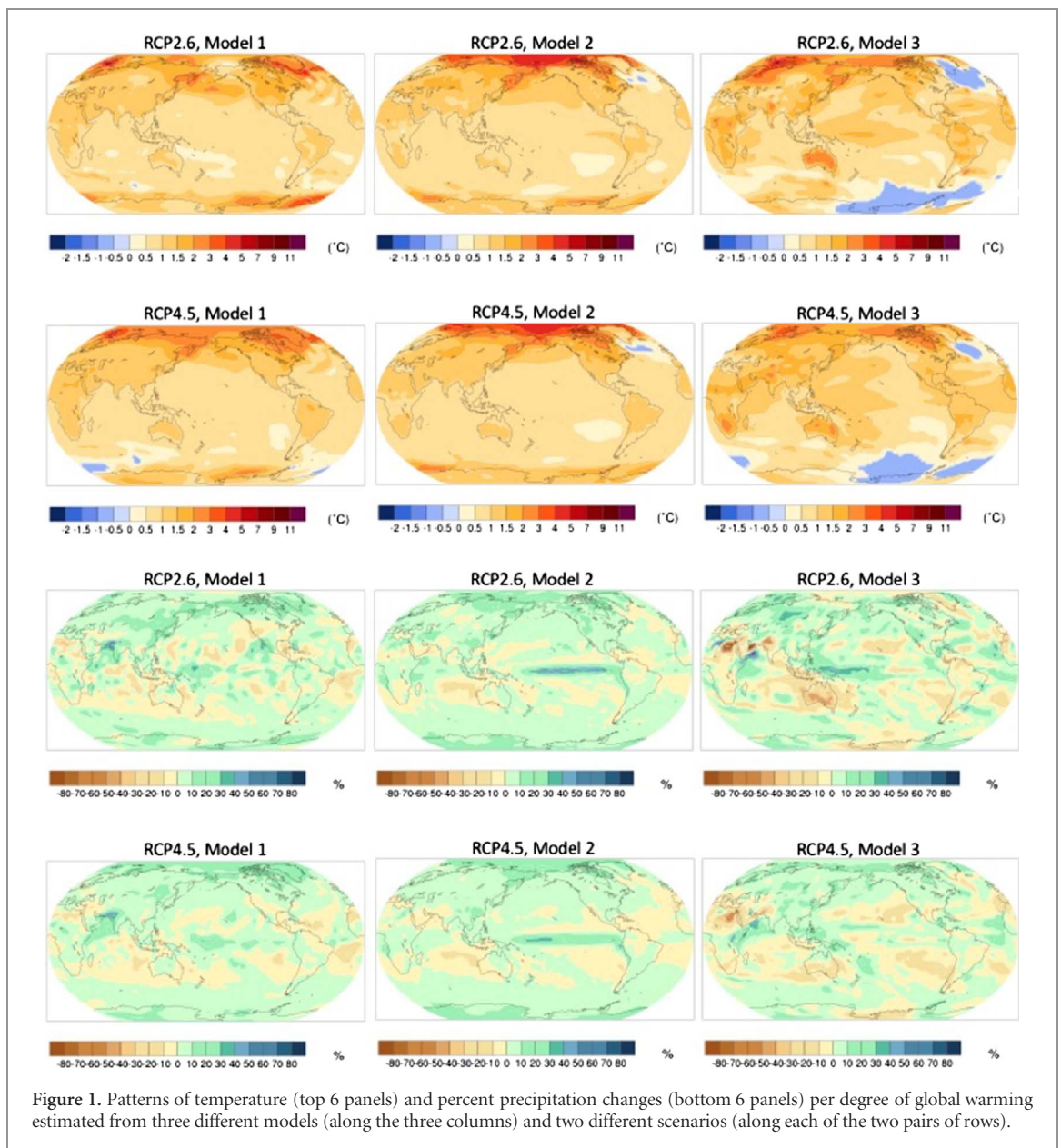
To evaluate the behavior of these annual fields, we compute their global averages and examine the characteristics of the variability of the time series, compared to the target, in terms of the magnitude of the interannual standard deviation and the presence or absence of a trend. Visual inspection of the time series characteristics is supplemented by quantitative assessment of the size of the interannual variability (true and emulated) through histograms and scatterplots.

Results

Emulation of RCP2.6

We use the multi-model spread as the standard against which to measure the magnitude of the emulation errors. Thus, we assess if the error introduced by the approximation is significant compared to the currently unavoidable uncertainty in the true target patterns attributable to the CMIP5 models' structural differences.

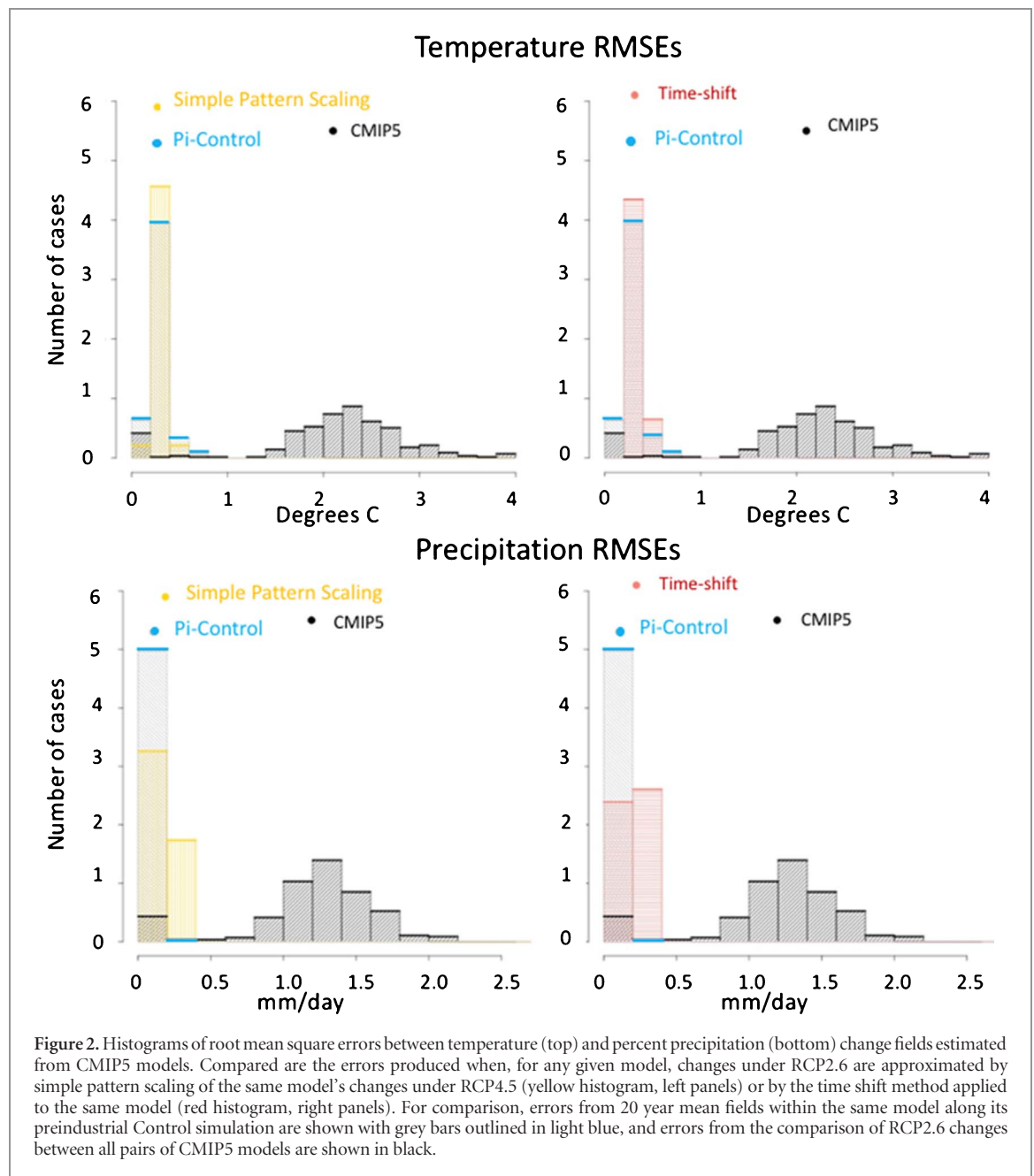
The top half of figure 1 for temperature change patterns, and its bottom half for percent precipitation patterns show that the variation in geographic features along rows (same scenario, three different models) is much larger than the variation along columns (same model, two scenarios). We therefore expect that errors from approximating scenarios *within the same model* will be of smaller magnitude than the variation of outcomes *across the multi-model ensemble for a given*



scenario. This is likely going to be the case when keeping a global perspective and summarizing errors across the spatial domain, rather than focusing on specific, regional, fine scale feature. Previous work has shown that there exist indeed preferential errors introduced by approximations in certain regions. For example, errors have been shown to be relatively larger over regions sensitive to polar amplification, i.e. located at high latitudes and therefore sensitive to the positive warming feedback induced by melting ice caps. Other types of local feedbacks may amplify the response to forcing in a non-linear way. Tebaldi and Arblaster 2014, Seneviratne *et al* 2006.

Note that all the models and simulations considered here are concentration driven, use fixed ice-sheets and prescribed land use and land cover changes. On the one hand, these aspects reduce the potential for differences in patterns across models and scenarios. On the other hand, they may be the source of inaccurate

representation of the spatial response. Expectations for how these different feedbacks may or may not affect the methods' performance are as follows. There is no obvious way by which the carbon cycle could change the pattern of warming (as opposed to the global mean), as CO₂ is well-mixed. Ice sheet feedbacks are unlikely to be large by 2100, since the volume of ice may change but the area of the ice sheets will hardly change by 2100, so the largest effect should be on sea level, with no effect on the pattern of warming. For vegetation, the forcing by land use and land use change is prescribed by the individual RCPs' specifications in the models, and interact in terms of surface feedbacks from snow and ice and other land surface characteristics. Questions arise only for interactive vegetation that is missing from the models. Only if interactive vegetation changed the pattern differently in different models, or differently across scenarios, then that would lead to poorer performance of the presented methods.



Which of the two methods presented would be more strongly affected is not clear, as vegetation feedbacks may depend both on the rate and on the magnitude of warming. However, such potential nonlinear changes (e.g. an Amazon dieback) are most relevant for very high warming levels, and since we are considering pattern emulations for low levels of warming we expect these effects to be very likely small. Specifically, we expect them to be small relative to the uncertainty that arises through the large differences in simulated physical feedbacks across models.

The visual impression from figure 1 is confirmed by the quantitative performance metric, RMSEs. Figure 2 presents histograms of RMSE values collected from the 23 models: the error introduced by simple pattern scaling when approximating the RCP2.6 response of a model by rescaling its response in RCP4.5 (left

panels, yellow histogram) or by using the time-shift approach within each model (right panels, red histogram) is here compared to two other histograms. The black bar plots in each panel, with a wide range covering larger values compared to the other histograms, show RMSEs from all possible pairs of RCP 2.6 patterns from the CMIP5 ensemble, thus measuring structural model uncertainty. The narrower grey histograms outlined in light blue, almost exactly overlaying the yellow and red bars, show the RMSEs obtained from differences in piControl patterns, giving an indication of the uncertainty introduced by internal variability. The significantly narrower and lower ranges of the histograms of emulation errors (red and yellow) compared to the histograms measuring inter-model variability (black) suggest that the inaccuracy of the approximation pales in comparison to the differences introduced in the

response to forcing by the models' structural uncertainty. At the same time, the similarity between the former and the histograms representing errors from internal variability suggests that the emulation inaccuracy is rather similar to the error due to the imperfect sampling of the forced component in any single model run. The comparison of RMSEs from precipitation emulation delivers very similar results. It has been established by recent work that 20 year averages at individual grid-points from a single model run can only start to identify the forced component in temperature changes (Deser *et al* 2012). Uncertainties only grow for other variables.

Next, we address the need of providing yearly output in the form of time series of temperature and precipitation changes for impact modeling, rather than 20 year averages. Scaling the 20 years individually would result in scaled (and therefore biased) internal variability. Some recent work has proposed sophisticated emulations of spatial variability in the context of pattern scaling (Alexeeff *et al* 2018). Here we simply reconstruct a 20 year time series taking the higher scenarios sequence of years and swapping out and in the forced component estimated by the pattern scaling method. Alternatively, we simply consider the 20 year window around the time-shifted target, as explained in the methods section. Rather than embarking on a complex evaluation and comparison of spatial variability of annual patterns between emulated and target scenarios, we simply ask how the globally averaged emulated annual patterns compare to the globally averaged target annual patterns. Figure 3 shows that the behavior of interannual variability emulated around pattern scaling is very similar to the behavior of true interannual variability, as inspection by eye of the time series of global means suggests. A quantitative comparison of the size of interannual standard deviations (collecting values from the different CMIP5 models) through histograms and scatter plots confirm the qualitative assessment. These results are to be expected, as many studies have shown that it is rare to detect significant and pervasive changes in variability for different levels of warming, even under high emission scenarios.

When we evaluate the time-shift approach the stronger trend present in the transient scenario (which at the time the 1.5 and 2.0 °C warming levels are reached is still subject to increasing CO₂ concentrations) and absent in the stabilized target scenarios introduces a spurious behavior in the emulated time series, enhancing interannual variability. This is true for temperature but much less for precipitation, whose trend in annual means under RCP4.5 is not significant. The removal of a simple linear fit in the case of temperature corrects the interannual variability of the emulation and draws it better in line with the target. In the case of precipitation, a trend removal appears to overcompensate.

Even if not the focus of this analysis, we call attention to the wide range of inter-annual variability

values that is present across the multi-model ensemble, as indicated by the range along the *x*-axis under the histograms. Models differ by factors of 2, 3, even 4 in the size of the simulated inter-annual fluctuations of global mean values of temperature and precipitation. Any study that depends on using this type of time series as input for climate impact models should carefully evaluate the climate models' internal variability, as the biases in the magnitude and patterns of variability are evidently significant. For example, we expect changes in indices defined as exceedances of absolute or relative thresholds, like spells of hot extremes, to be particularly sensitive to the size of the model internal variability.

Emulation of 1.5 °C and 2.0 °C scenarios

In addition to exploring the performance of these two emulation techniques in the multi-model framework of CMIP5, where our target was RCP2.6, we provide results of the same emulation methods applied in the single-model, initial condition ensemble framework, where we can target specifically the two scenarios motivated by the Paris agreement.

For this application of the methods the standard will be the distribution of RMSEs expected solely because of internal variability. Specifically, we compute these RMSEs by considering differences between every possible pair of target patterns obtained from the 10 initial condition ensemble members available under each scenario.

We summarize results in the form of RMSEs histograms as before. Figure 4 demonstrates that even when focusing on the single-model behavior the two emulation techniques (yellow and red histograms) produce, across the board, RMSEs of the same range of magnitudes as the presence of internal variability does (grey histograms). This is the case for both quantities (temperature and precipitation), both scenarios (1.5 °C and 2.0 °C), both methods (simple pattern scaling and time-shift). We show maps of the true and error fields for temperature and precipitation, the two scenarios and the two emulations in the supplementary material (figures S1–S6 available at stacks.iop.org/ERL/13/055006/mmedia). The similarity of the target and emulated patterns and the absence of systematic error patterns (with the possible exception of a wet bias for the emulation of precipitation changes in the high variability region of the western Pacific region for the 1.5 °C scenario) can be appreciated from these maps.

Also the analysis of inter-annual variability produces similar results to what we described for RCP2.6, as figure 5 shows: pattern scaling performs accurately in this exercise, while the time series of annual quantities identified through the time-shift method are affected by the presence of an overall trend that needs correcting, in the case of these scenarios more substantially than before. Here, since we are considering a single model, the size of interannual variability is not significantly

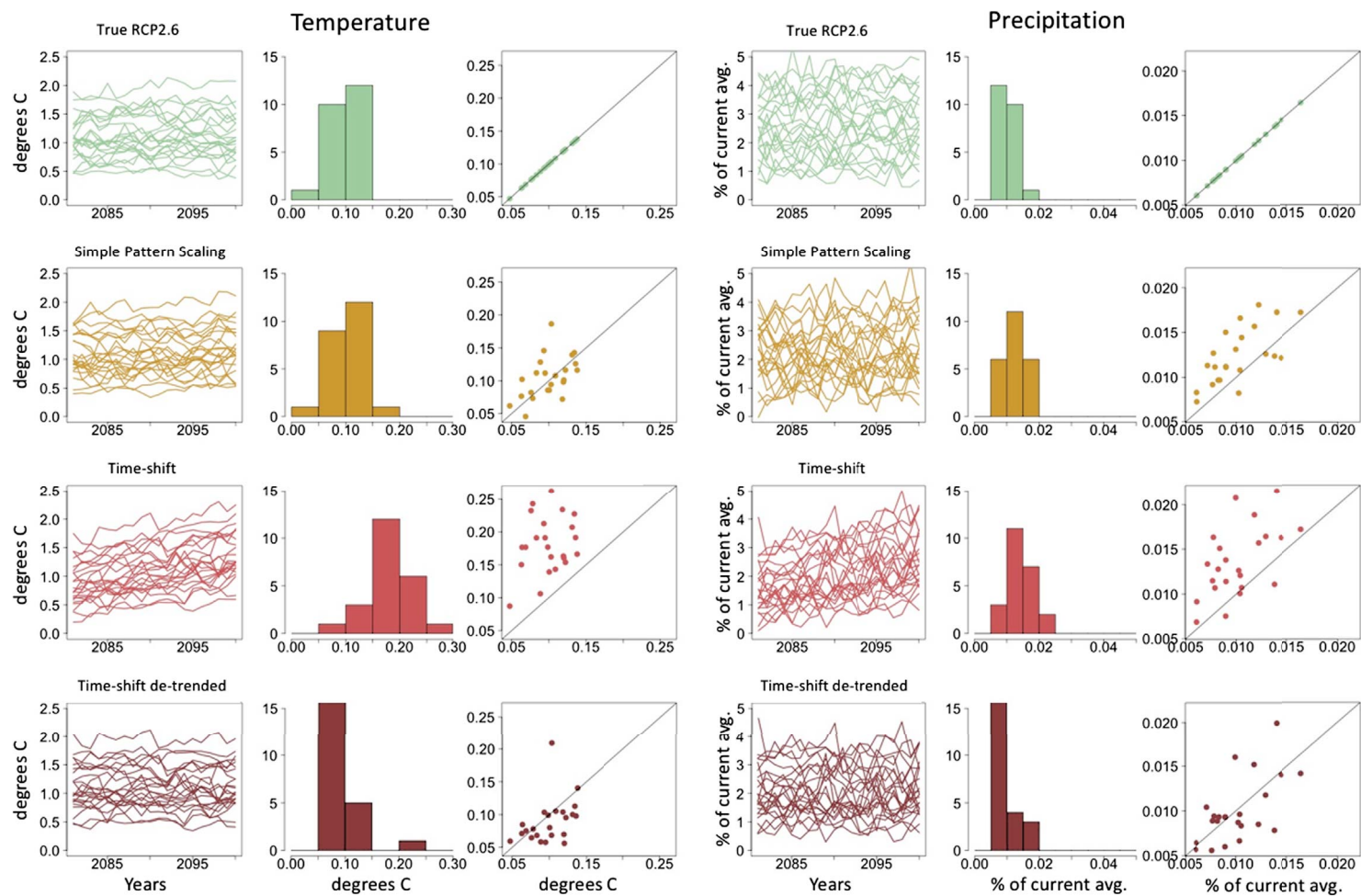


Figure 3. Comparison of interannual variability of global means of temperature and percent precipitation change for CMIP5 during the last 20 years of the 21st century under RCP2.6 and by approximating them by pattern scaling and time shift (see plot titles). Time series are shown along the left columns of the two sets, while histograms of the interannual variability across models are shown in the middle column (with y-axis showing the number of cases (models)). The last histogram shows the interannual variability from the time shift approximation after linearly detrending the trajectories. Scatter plots along the right columns compare true values and values from approximations of the interannual variability. Units are the same on both axes for the scatter plots.

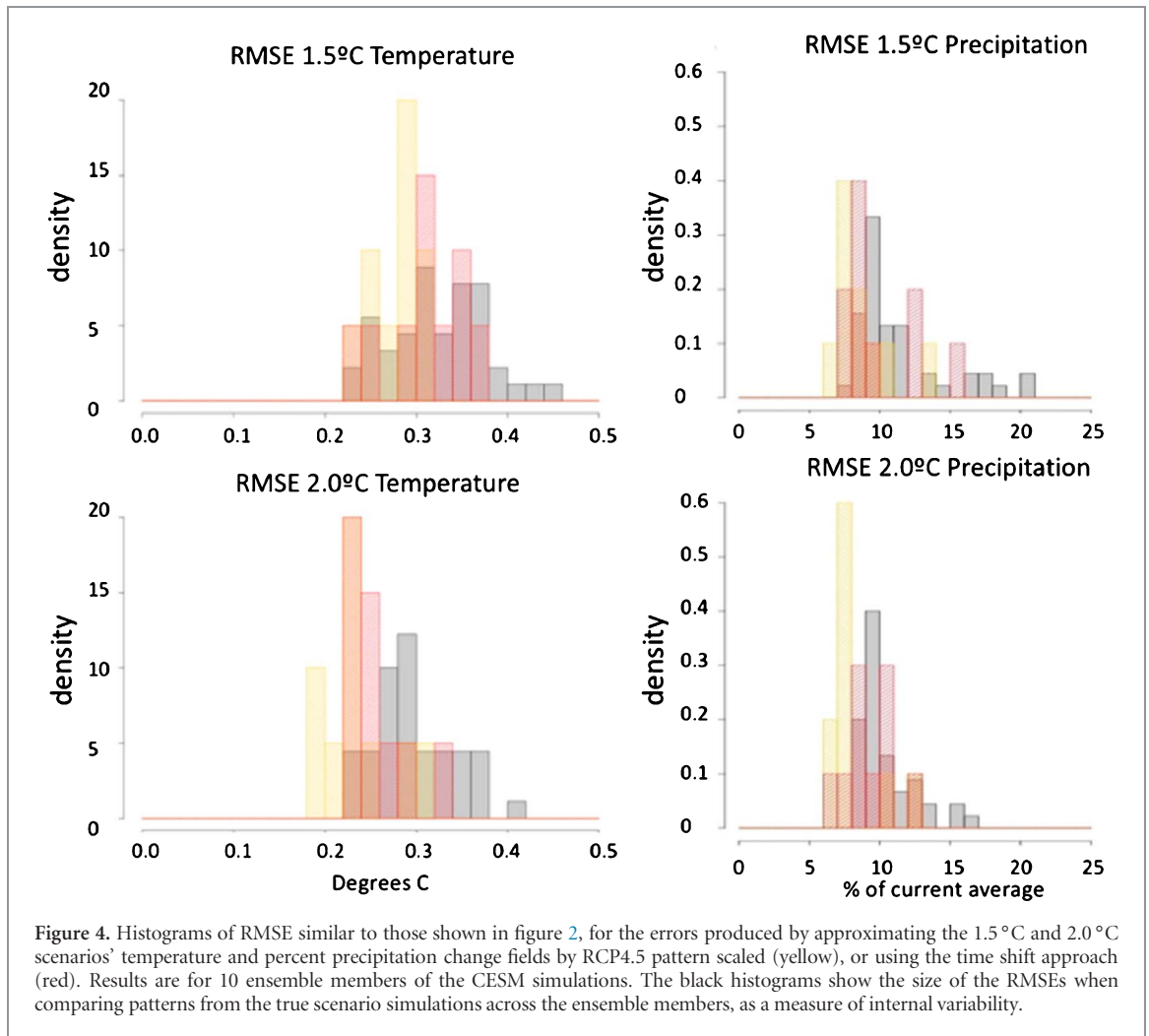


Figure 4. Histograms of RMSE similar to those shown in figure 2, for the errors produced by approximating the 1.5 °C and 2.0 °C scenarios' temperature and percent precipitation change fields by RCP4.5 pattern scaled (yellow), or using the time shift approach (red). Results are for 10 ensemble members of the CESM simulations. The black histograms show the size of the RMSEs when comparing patterns from the true scenario simulations across the ensemble members, as a measure of internal variability.

different across the ensemble members, as opposed to what was noted for the multi-model ensemble results. The question however remains whether the interannual variability simulated by CESM1 compares favorably to the observed, and, if a discrepancy is assessed, what the impact model's tolerance is, in order to still produce meaningful projections, once other sources of uncertainties are accounted for.

Discussion and conclusions

This study was motivated by the need for evaluating the differential impacts of alternative low warming scenarios in the wake of the Paris agreement. We take advantage of the availability of simulations conducted with NCAR-DOE CESM1 to test the accuracy of emulation approaches that can approximate such scenarios based on available CMIP5 type scenarios. In fact, we cast the net wide at first, asking if the highly mitigated RCP2.6 can be approximated by RCP4.5 in the context of a multi-model ensemble. In that case, the accuracy of the emulation is judged by comparing emulation errors to the uncertainty introduced by the diversity of responses to the same forcing pathway that the multi-model ensemble produces.

We investigate the accuracy of two emulation methods, simple pattern scaling and time-shift, for two target quantities: annual average temperature and precipitation. We find that both methods produce emulated patterns of temperature and precipitation change that approximate the forced response (a 20 year average change by the end of the century) well within the CMIP5 models' structural uncertainty. The same is true even when comparing approximation errors to the errors introduced by the presence of model internal variability. After we reconstruct inter-annual variability and superimpose it on the forced pattern, we find that doing so on the basis of pattern scaling performs better than using the annual time series from the time-shift approach, because of the presence of a trend within the time window used in the RCP4.5 simulation to approximate the low warming scenarios. This could be remedied by subtracting a linear trend.

For impact analysis that can build on these basic quantities, this study implies that existing simulations under the CMIP5 protocol could be used to approximate the outcomes of the new low-warming scenarios, as most changes appear to be linear in global average temperature. For these type of applications, therefore, the lack of specific simulations approximating 1.5 °C

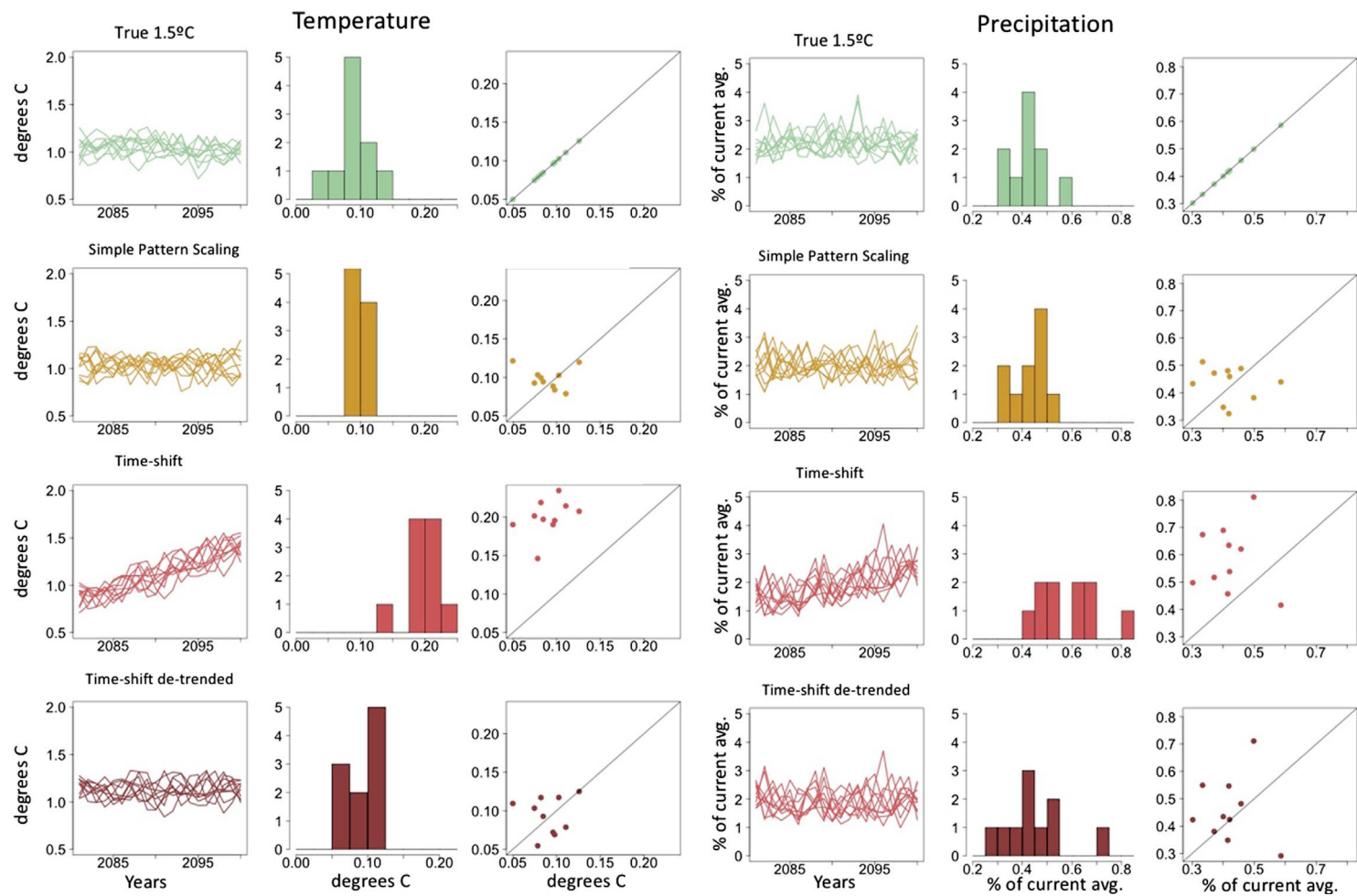


Figure 5. Comparison of interannual variability of global means of temperature and percent precipitation change from CESM during the last 20 years of the 21st century under the 1.5°C scenario and by approximating them by pattern scaling and time shift (see plot titles). Time series are shown along the left columns of the two sets, while histograms of the interannual variability across models are shown in the middle column (with y-axis showing the number of cases, i.e. ensemble members). The last histogram shows the interannual variability from the time shift approximation after linearly detrending the trajectories. Scatter plots along the right columns compare true and approximations. Units are the same on both axes for the scatter plots. Figure S7 shows analogous results for the emulation of the 2.0°C scenario.

and 2.0 °C by global climate models is unlikely to be a limiting factor, and approximations through scaling of climate model output should not make a substantial contribution to the overall uncertainty, starting from that introduced by considering multiple climate models.

While assessing the behavior of these emulation techniques, the reality of the large discrepancies of the multi-model ensembles (with regard to both patterns of change and the size of internal variability) surfaces starkly and calls for attention: model diversity continues to be necessary input to any robust impact analysis. Inter-annual variability should be easier to diagnose and evaluate on the basis of observations than patterns of change. We therefore suggest that if the impact model is sensitive to it, it should be beneficial to pay attention to this aspect and possibly select models on its basis.

This has been a rather basic exploration, and we are aware of other compelling needs when thinking of drivers of impacts, such as the joint behavior of multiple variables, or their tail behavior through metrics of extremes. Their emulation should be tackled and tested too, in order to provide a more complete suite of climate drivers to the impact research community in the quest for assessing the benefit of slight differences in global warming levels.

Acknowledgments

This project was supported by the Regional and Global Climate Modeling Program (RGCM) of the US Department of Energy, Office of Science (BER), Cooperative Agreement DE-FC2-97ER62402. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. For CMIP the US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support

and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. We are also grateful to Jan Sedláček for helping with the graphics.

ORCID iDs

Claudia Tebaldi  <https://orcid.org/0000-0001-9233-8903>

References

- Alexeeff S E *et al* 2018 Emulating mean patterns and variability of temperature across and within scenarios in anthropogenic climate change experiments *Clim. Change* **146** 319–33
- Deser C *et al* 2012 Uncertainty in climate change projections: the role of internal variability *Clim. Dyn.* **30** 527–46
- Herger N *et al* 2015 Improved pattern scaling approaches for the use in climate impact studies *Geophys. Res. Lett.* **42** 3486–94
- James R *et al* 2017 Characterizing half-a-degree difference: a review of methods for identifying regional climate responses to global warming targets *Wiley Interdiscip. Rev. Clim. Change* **8** e457
- King A *et al* 2017 Australian climate extremes at 1.5 °C and 2 °C of global warming *Nat. Clim. Change* **7** 412–6
- Meinshausen M *et al* 2011 Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6—Part 1: model description and calibration *Atmos. Chem. Phys.* **11** 1417–56
- Mitchell D *et al* 2017 Half a degree additional warming, prognosis and projected impacts (HAPPI): background and experimental design *Geosci. Model Dev.* **10** 571–83
- Moss R *et al* 2010 The next generation of scenarios for climate change research and assessment *Nature* **463** 747–56
- Sanderson B *et al* 2018 A new ensemble of GCM simulations to assess avoided impacts in a climate mitigation scenario *Clim. Change* **146** 303–18
- Sanderson B *et al* 2017 Community climate simulations to assess avoided impacts in 1.5 and 2 °C futures *Earth Syst. Dyn.* **8** 827–47
- Santer B *et al* 1990 Developing climate scenarios from equilibrium GCM results Max-Planck-Institut für Meteorologie Report
- Seneviratne S *et al* 2006 Land-Atmosphere coupling and climate change in Europe *Nature* **443** 205–9
- Taylor K E *et al* 2012 An overview of CMIP5 and the experiment design *Bull. Am. Meteorol. Soc.* **93** 485–98
- Tebaldi C and Arblaster J M 2014 Pattern scaling: its strengths and limitations, and an update on the latest model simulations *Clim. Change* **122** 459–71