Sensitivity of regional climate to global temperature and forcing

You may also be interested in:

Half of the world's population experience robust changes in the water cycle for a 2 degree C warmer world
Jan Sedláek and Reto Knutti

The physical drivers of historical and 21st century global precipitation changes
Livia Thorpe and Timothy Andrews

Rapid and extensive warming following cessation of solar radiation management
Kelly E McCusker, Kyle C Armour, Cecilia M Bitz et al.

Emergence of new hydrologic regimes of surface water resources in the conterminous United States under future warming
Guoyong Leng, Maoyi Huang, Nathalie Voisin et al.

Poorest countries experience earlier anthropogenic emergence of daily temperature extremes
Luke J Harrington, David J Frame, Erich M Fischer et al.

Assessing climate change impacts, benefits of mitigation, and uncertainties on major global forest regions under multiple socioeconomic and emissions scenarios
John B Kim, Erwan Monier, Brent Sohngen et al.

The response of the climate system to very high greenhouse gas emission scenarios
Benjamin M Sanderson, Brian C O'Neill, Jeffrey T Kiehl et al.

Carbon budgets and energy transition pathways
Detlef P van Vuuren, Heleen van Soest, Keywan Riahi et al.

Environmental Research Letters

Sensitivity of regional climate to global temperature and forcing

Claudia Tebaldi[1], Brian O'Neill[1] and Jean-François Lamarque[1,2]

[1] Climate and Global Dynamics Division, National Center for Atmospheric Research, PO BOX 3000, Boulder, CO, 80307, USA
[2] Atmospheric Chemistry Division, National Center for Atmospheric Research, PO BOX 3000, Boulder, CO, 80307, USA

E-mail: Tebaldi@ucar.edu

Abstract

The sensitivity of regional climate to global average radiative forcing and temperature change is important for setting global climate policy targets and designing scenarios. Setting effective policy targets requires an understanding of the consequences exceeding them, even by small amounts, and the effective design of sets of scenarios requires the knowledge of how different emissions, concentrations, or forcing need to be in order to produce substantial differences in climate outcomes. Using an extensive database of climate model simulations, we quantify how differences in global average quantities relate to differences in both the spatial extent and magnitude of climate outcomes at regional (250–1250 km) scales. We show that differences of about 0.3 °C in global average temperature are required to generate statistically significant changes in regional annual average temperature over more than half of the Earth's land surface. A global difference of 0.8 °C is necessary to produce regional warming over half the land surface that is not only significant but reaches at least 1 °C. As much as 2.5 to 3 °C is required for a statistically significant change in regional annual average precipitation that is equally pervasive. Global average temperature change provides a better metric than radiative forcing for indicating differences in regional climate outcomes due to the path dependency of the effects of radiative forcing. For example, a difference in radiative forcing of 0.5 W m^{-2} can produce statistically significant differences in regional temperature over an area that ranges between 30% and 85% of the land surface, depending on the forcing pathway.

Introduction

Future pathways or targets expressed in terms of global quantities such as global average surface temperature change (GAT), radiative forcing (RF) or atmospheric concentrations are used to define future scenarios and play an important role in climate policy and science. Most prominently, the only internationally agreed upon policy goal is to limit the increase of GAT to 2 °C above pre-industrial levels (UNFCCC 2009). Other global quantity goals have served to structure policy discussions, and scientific analysis and assessment have elucidated conditions associated with particular goals, including mitigation costs, impacts and adaptation options (Meinshausen *et al* 2009, Huntingford *et al* 2012, Schaeffer *et al* 2012, Frieler *et al* 2013, Manoj *et al* 2011, Rogelj *et al* 2011, 2014, Sedlacek and Knutti 2014, Oppemheimer *et al* 2014, Clarke *et al* 2014, NRC 2011).

These global metrics of future changes are associated with regional changes in climate that are directly responsible for impacts, and analyses have focused on the consequences of different global targets, such as 450 versus 550 ppm CO_2 equivalent concentrations (Clarke *et al* 2007, Waldhoff *et al* 2014) or 2 °C versus 4 °C (New *et al* 2011). However, little analysis has been devoted to understanding the sensitivity of regional climate outcomes to marginal variations in global targets (e.g., exceeding 2 °C by a few tenths of a degree C). Marginal differences are important for understanding what the impact consequences might be of exceeding a given global target, including by temporarily (and possibly intentionally) overshooting it and returning to it later, whether through mitigation or geoengineering (Lowe *et al* 2009, Wigley 2006).

Anticipated marginal differences in regional climate outcomes also play a central role in the choice of

future scenarios to run in large, resource-intensive climate model comparison exercises. In a previous exercise (Coupled Model Intercomparison Project Phase 5, CMIP5; Taylor *et al* 2012), four scenarios were chosen in part based on the undocumented assumption that a separation in global radiative forcing levels in 2100 of approximately 2 Wm^{-2} was required for significant differences in outcomes (Moss *et al* 2010). A new set of scenarios is currently being chosen for a new comparison exercise (CMIP6; Meehl *et al* 2014, van Vuuren *et al* 2014), and a better-grounded criterion for forcing separation is essential to this process.

In this study we quantify the sensitivity of regional climate outcomes to global quantities by drawing on results of both idealized experiments and realistic forcing scenarios from up to 29 climate models from the CMIP5 database. For a given change in a global metric like GAT or RF we measure the differences in average annual temperature and precipitation for individual grid cells (~2.5 degrees, or ~250 km at the equator), assess their statistical significance (at the 5% level using a null distribution derived from pre-industrial control runs), and summarize the pervasiveness of these changes as the fraction of the Earth's surface (or land surface only) significantly affected. Repeating this for small increments in the global metrics, we derive an empirical relation between the size of the global scenario differences and the size and significance of the regional (impact-relevant) differences associated with them. By using the CMIP5 framework we account for model uncertainty; in addition we test the sensitivity of our results to the use of coarser spatial scales and seasonal rather than annual averages.

Recent studies have explored related questions, in particular characterizing the ratios of forced signal to internal variability and the times of emergence of the forced response under future forcings (Hawkins and Sutton 2009, 2012, Baettig *et al* 2007, Deser *et al* 2012, Giorgi and Bi 2009, Tebaldi and Friedlingstein 2013, Mahlstein *et al* 2012), within a regional perspective. However these studies have not focused on the question of differences between scenarios, nor systematically quantified these differences as a function of global quantities as we do in this study.

In the next section, we present the methodology in more detail. We then discuss results and their robustness to a number of methodological choices. The last section concludes, discussing implications as well as possible extensions and future work.

Methods

We draw on climate model simulations for 5 different scenarios: an idealized 1% CO_2 increase per year experiment, a 1% CO_2 increase per year followed by stabilization at 1 or 2 W m^{-2}, historical (1850–2005) emissions of multiple gases and aerosols and land use change, and the RCP4.5 scenario of future emissions and land use change that stabilizes at 4.5 W m^{-2}.

We use 29 models from the CMIP5 archive for the results in the main text from the 1% CO_2 increase per year experiments, and a subset of 19 when comparing to historical and RCP4.5 results. These 19 models are the ones that ran RCP4.5 (27 of them are available) and also reached 2 °C of ΔGAT within that experiment (which eliminates 8 of these), a necessary condition in order to perform the sensitivity analysis to different baseline warming levels (see results section). The list of models is available as supplementary material (available at stacks.iop.org/ERL/10/074001/mmedia). The results described for the stabilization experiments at 1 or 2 W m^{-2} are based on the CESM1 model, since stabilized experiments of this type are not available as part of the CMIP5 experimental design.

We use one ensemble member for each model. The analysis is developed for each model separately (i.e. based on a model's trajectory of global average temperature change or associated trajectory of RF levels and corresponding fields of temperature or precipitation changes at the grid-point scale). We first regrid all models to a common T42 Gaussian grid (with a grid spacing of ~2.5 degrees, i.e., ~250 km at the equator). We then compute the individual model results in terms of significance and magnitude of the regional differences, then aggregate the results through multi-model summaries (medians and boxplots).

Significance of the changes is computed separately for the results from each model, with reference to a null distribution of changes derived from the pre-industrial control runs available for the individual model. The null distribution is derived at each grid point for both annual and seasonal average surface temperature (TAS) and precipitation (PR) by computing a large set of twenty-year averages from non-overlapping segments of the control run. We calculate differences between all possible pairs of averages, thus deriving a distribution of differences that we consider the reference distribution against which to evaluate the significance of the twenty-year differences that we compute from the scenario runs of the same model (1% CO_2 increase per year, transient and stabilized, historical and RCP4.5).

To compute the differences in the scenario runs, we consider either the annual time series of GAT for each individual model integration, or a common time series of RF levels that can be associated with all models' integrations in the case of 1% CO_2 increase per year experiments, or the time series of RF levels associated with the CESM1 stabilization experiments. We choose as a baseline the average of the first 20 years of the 1% CO_2 increase per year integrations for both GAT and RF or the first 20 years of the other scenario experiments. When we test the robustness of results to higher temperature baselines we use the twenty-year average temperature values at the time when GAT is 1 or 2 °C above those same initial averages.

For each experiment, the corresponding time series of annual GAT or RF is first smoothed with a twenty-year running average, and then used to determine the point in time when a given change in those global quantities is reached.

For the same time, the corresponding twenty-year average change in surface temperature or precipitation from the baseline is computed at the grid-cell scale and its significance tested with respect to the reference distribution from the control run. For TAS, we consider a difference to be significant if its value falls in the tail of the null distribution to the right of the 95th quantile (akin to a one-sided *t*-test at the 5% level). For PR, we consider a difference significant if its value falls in the tails of the null distribution to the left of the 2.5th quantile or to the right of the 97.5th quantile (akin to a two-sided *t*-test at the 5% level).

Results are aggregated for each model run by measuring the fraction of the Earth's surface (or land surface only) over which a significant change takes place, weighing each grid point by its relative area (the cosine of its latitude). When multi-model results are shown as maps or through distributions of the magnitude of regional differences (e.g. figure 4), we first determine at which locations at least half of the models available show a significant change, then show the median value of surface temperature change or precipitation change from those models producing a significant change.

For the main results in this paper, as explained, regional climate outcomes are defined as differences in average annual temperature and precipitation for individual grid cells (∼2.5 degrees, or ∼250 km at the equator), but we also test the sensitivity of results to the use of seasonal means and to coarser spatial scales. We assess both the statistical significance and the absolute magnitude of these differences as a function of incremental changes in global radiative forcing (ΔRF) and global average temperature increase (ΔGAT).

Results

Sensitivity of regional climate

Summary maps of multi-model ensemble outcomes for regional annual temperature change above the baseline period (the first 20 years of the simulations) derived from the 1% CO_2 increase per year experiments show that, as expected, the warming signal is larger and more widespread for larger values of ΔGAT, that it emerges first in the tropics where natural variability is smallest (Mahlstein *et al* 2011), and that it is most pronounced at high latitudes (figure 1). Maps of outcomes for regional annual precipitation change show different patterns (figure 2): the signal emerges first at high latitudes, and eventually emerges in the equatorial pacific and in the regions of the mid-to-low latitudes already identified as prone to drying, all of which is consistent with the already well-documented response of the hydrological cycle to warming (Held

and Soden 2006). For both variables the regional signal is larger for larger values of ΔGAT, but regional outcomes for temperature are much more sensitive to global temperature or forcing than are outcomes for precipitation. Strikingly, a ΔGAT of 1 °C creates pervasive significant changes in mean annual temperatures but produces significant changes in annual precipitation over only a small fraction of the surface.

Summary distributions of individual models' behavior (figure 3) show that the spatial extent of the effect on regional climate increases nonlinearly with changes in global quantities. We focus on outcomes for the land surface, which we take to be more relevant to a range of environmental impacts than results for the entire surface (see figure S1 in the supplementary material for results for the entire Earth's surface, for which all our conclusions remain valid). For example, in order for a majority of models to show a significant change over at least half of the land surface, ΔGAT must increase by about 0.3 °C (top left panel), and ΔRF by about 0.75 W m^{-2} (bottom left panel). However, significant change for the entire land surface in most models has not occurred completely even at 1 °C of ΔGAT and 2 W m^{-2} ΔRF.
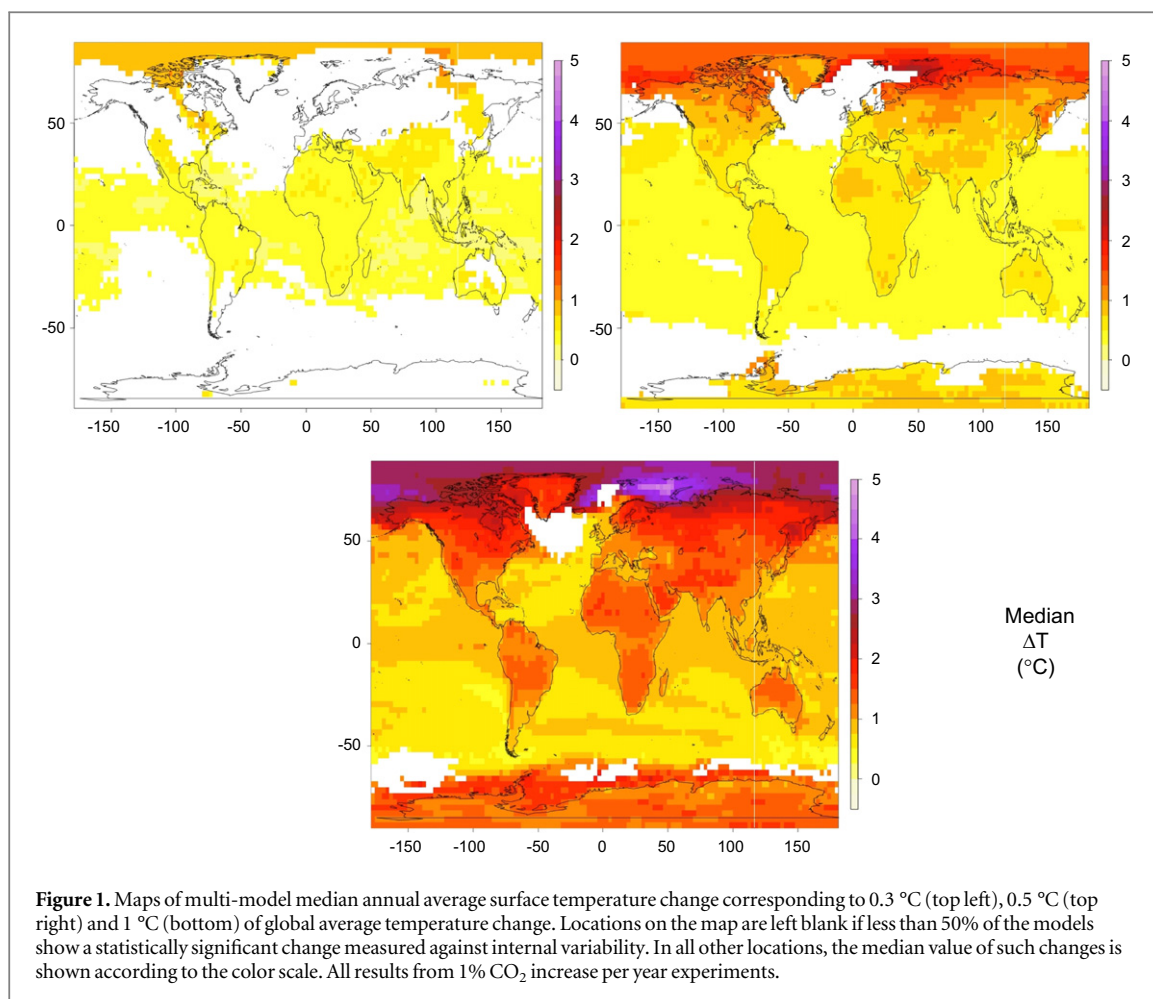
Note that the separation rule of 2 W m^{-2} de facto applied to the choice of RCPs for CMIP5 produces significant temperature changes over at least 90% of the land surface. Arguably, many impact analyses could detect important consequences at much lower thresholds.

A comparison of the results based on ΔGAT and ΔRF highlights an important, if perhaps expected, point: conditioning the analysis on ΔGAT reduces the uncertainty in the results, since it controls for inter-model differences in climate sensitivity. Any given value of ΔRF is associated with different levels of warming across models due to their different climate sensitivities, producing larger differences in significance levels. As we will discuss below, results based on ΔRF also turn out to be more sensitive to the shape of the scenario used in the analysis.

Figure S2 shows similar boxplots for seasonal changes in temperature for December–January–February and June–July–August means, and both land-only and the entire surface. As expected, seasonal results are slightly different than annual results as a consequence of the different variability affecting the seasonal mean quantities, but a difference in GAT of at most 0.4 °C is sufficient to produce significant differences in seasonal mean temperature over at least half the land surface (as opposed to 0.3 °C for annual means).

Results show that regional precipitation change is also related nonlinearly to global average temperature change (figure 3), but that the statistically significant spatial extent of the effect is much smaller than in the case of regional temperature outcomes. Pervasive, statistically significant precipitation changes are not achieved within the individual models even for ΔGAT

Figure 1. Maps of multi-model median annual average surface temperature change corresponding to 0.3 °C (top left), 0.5 °C (top right) and 1 °C (bottom) of global average temperature change. Locations on the map are left blank if less than 50% of the models show a statistically significant change measured against internal variability. In all other locations, the median value of such changes is shown according to the color scale. All results from 1% CO_2 increase per year experiments.

of 4 °C or ΔRF of 4 W m^{-2}. For ΔGAT of about 2.5 °C, however, the majority of models show significant precipitation changes over at least 50% of the land surface.

Figure S3 shows results for seasonal changes. Here the criterion to achieve the same significance across the Earth's surface as for annual averages is shown to change by about 1 °C (or more in the case of June–July–August (JJA) results for land-only), as could be anticipated by the more volatile nature of precipitation means: warming of GAT substantially larger than 3 °C is required for the majority of models to show significant precipitation changes over at least 50% of the surface.
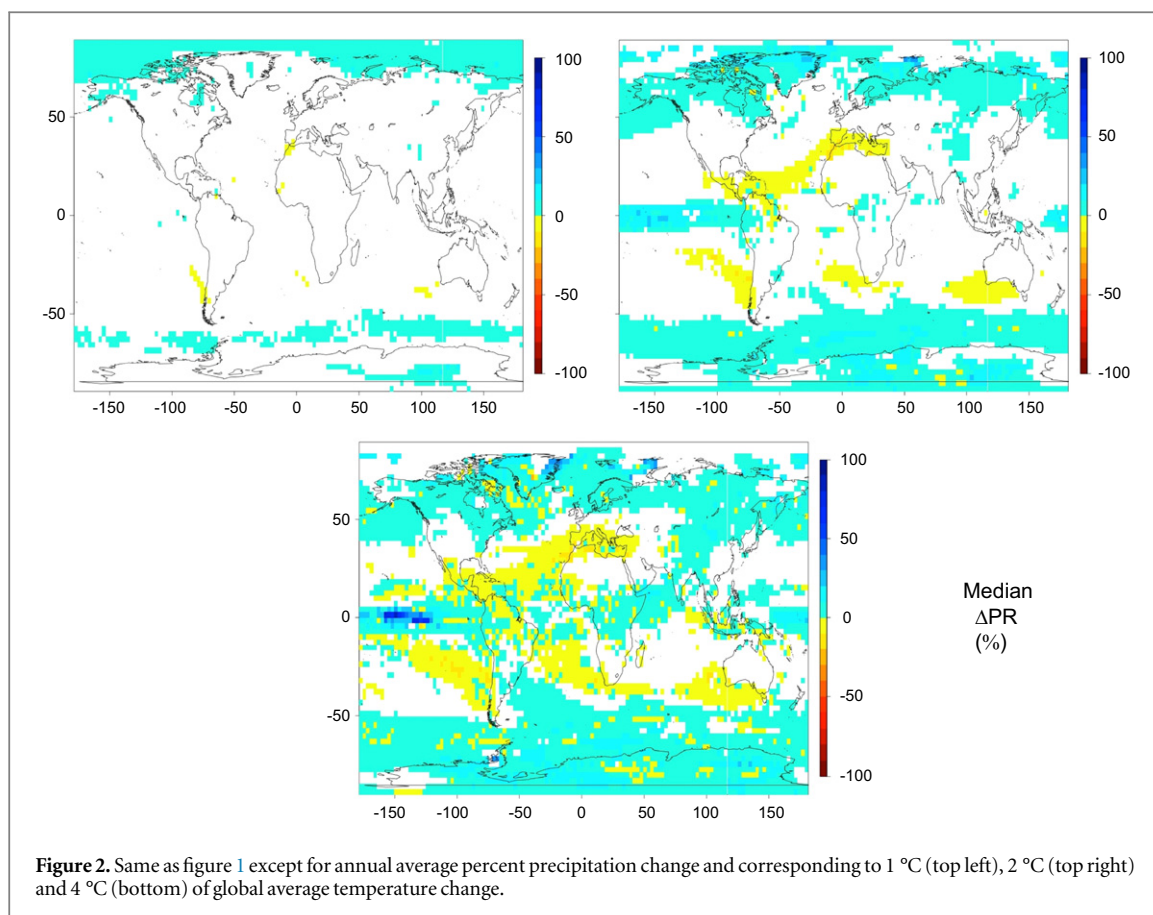
In many impact or adaptation studies it is the magnitude of regional climate change to which outcomes are most sensitive. Figure 4 adds magnitude to extent in multi-model summary distributions of results, showing that applying a criterion in terms of magnitude of warming can substantially increase the required difference in global quantities. For example, if we impose the criterion that at least half of the land area show a significant warming of *at least* 1 °C (by at least half of the models, measured as their median change), then a ΔGAT of 0.8 °C is required (left panel). Compare this to the required 0.3 °C ΔGAT when only statistical significance of the change is used as a criterion, regardless of magnitude. Figure S4 (left panels)

shows that the same required difference in GAT would apply to seasonal results.

The magnitude of regional precipitation changes, in absolute values, remains between 10% and 30% of the baseline average precipitation for the majority of the land affected, with only 10% of the surface affected by changes of 40% or more if warming of global temperature exceeds about 3.5 °C. Figure S4 (right panels) provides seasonal results, for which changes in regional precipitation larger than 50% can be expected for the same large levels of global warming, especially in December–January–February (DJF).

Figures S5 and S6 of the supplementary materials provide annual average results for land and ocean area combined and a breakdown of the precipitation results into positive and negative changes.

Contrasting the results for precipitation and temperature highlights that regional temperature changes are likely easier to detect than most other climate outcomes (which have larger internal variability). Consequently the values of ΔGAT needed to achieve a given extent of significant regional outcomes for most other climate variables will be at least as large as those derived for regional temperature change outcomes, with the possible exception of some types of extremes whose changes are expected to be larger than the changes in mean.

Figure 2. Same as figure 1 except for annual average percent precipitation change and corresponding to 1 °C (top left), 2 °C (top right) and 4 °C (bottom) of global average temperature change.

Robustness of results

We tested the robustness of our results in several respects: averaging regional outcomes over a larger surface area; the use of a realistic scenario with multiple forcings rather than the idealized 1% CO_2 increase per year; the use of a baseline in which some warming has already taken place; and the use of a radiative forcing stabilization pathway rather than a continuous increase in forcing to test the path dependency of changes.
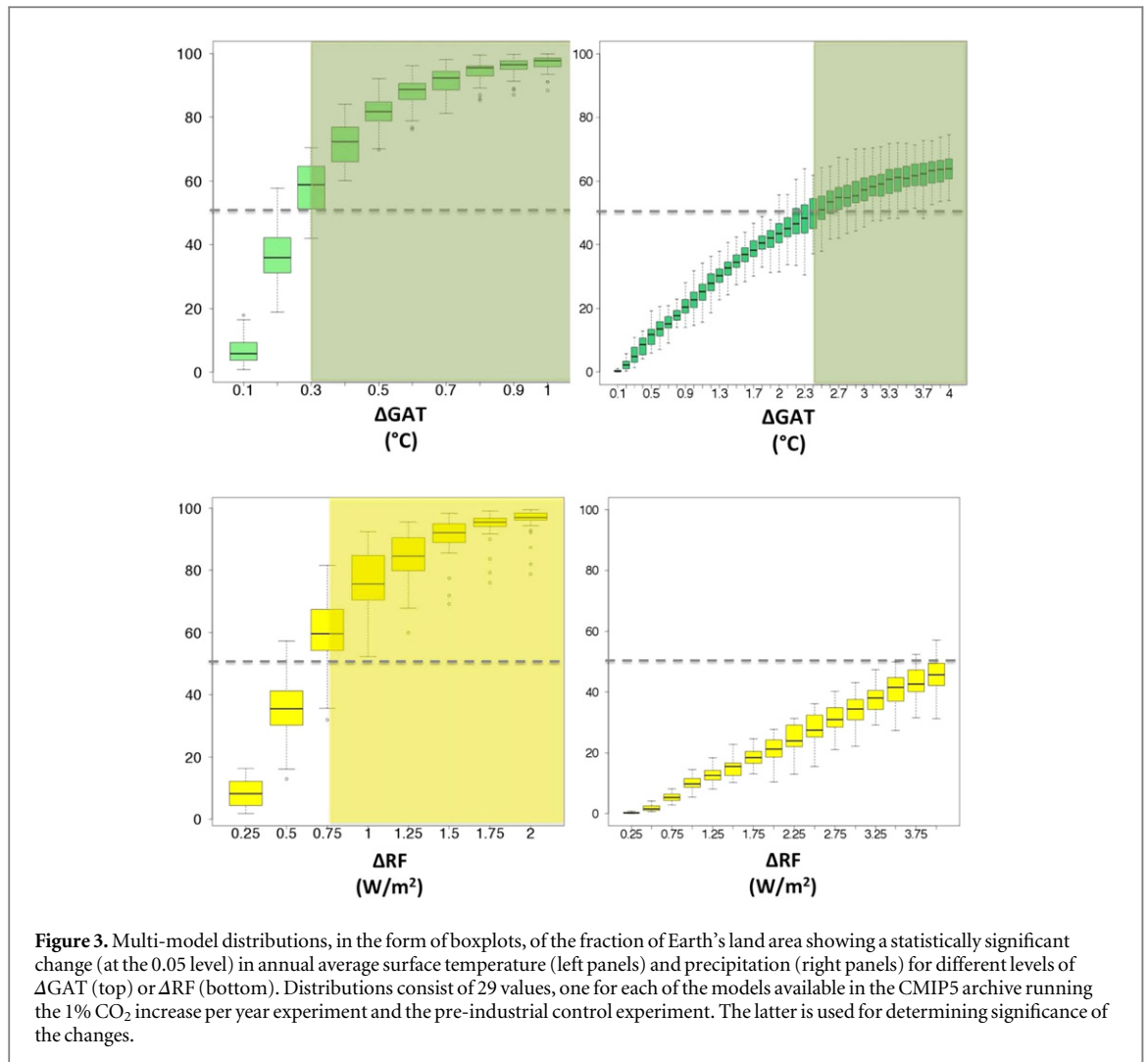
Results were robust to averaging temperature and precipitation fields over two increasingly large areas (9 and 25 grid cells, or about 750–1250 km at the equator; see figures S7 and S8 of the supplementary material). As expected, for a given level of ΔGAT, spatially aggregated output generates a larger spatial extent of significant regional change, since the spatial averaging reduces the noise from internal variability and therefore allows the signal to emerge more quickly. The thresholds identified earlier however remain valid for the two coarser scales considered: in particular, for temperature, 0.3 °C is still required to achieve significant changes over at least half of the land surface, and 0.8 °C produces a warming of at least 1 °C over that same portion.

To test the robustness of our results to the use of realistic scenarios we used historical experiments in which all forcings (greenhouse gases, aerosols, land use) change over the length of the simulation, spanning the years 1850–2005. The relationships between ΔGAT or ΔRF and regional outcomes were largely similar to the results based on the 1% CO_2 increase per year simulations (see figure S9 for boxplot summaries); regionally specific outcomes differed only over some regions of the Northern Hemisphere, where changes were less homogeneously significant (see figure S10, top panels comparing 1% CO_2 increase per year run to historical run results). This result would seem consistent with the cooling effects of aerosol precursor emissions in the historical run.

To investigate results with respect to a baseline in which 1 or 2 °C of warming have already occurred, we used the historical period combined with RCP4.5. The relationships between ΔGAT and regional outcomes (figure S10, bottom panels) are similar to those derived from the 1% CO_2 increase per year runs in terms of the pattern and extent of significant changes, with the exception of a smaller extent of statistically significant temperature differences over Northern Europe when the baseline is +2 °C. This outcome is consistent with a slowdown of the Atlantic Meridional Overturning Circulation, which models simulate at about that level of warming (+2.5 °C with respect to pre-industrial; Collins *et al* 2013).

Lastly, based on experiments conducted with the NCAR-DOE CESM1 model (Hurrell *et al* 2013, Meehl *et al* 2013), we tested the sensitivity of the relationship between ΔRF and regional outcomes to the forcing pathway by analyzing two scenarios in which CO_2

Figure 3. Multi-model distributions, in the form of boxplots, of the fraction of Earth's land area showing a statistically significant change (at the 0.05 level) in annual average surface temperature (left panels) and precipitation (right panels) for different levels of ΔGAT (top) or ΔRF (bottom). Distributions consist of 29 values, one for each of the models available in the CMIP5 archive running the 1% CO_2 increase per year experiment and the pre-industrial control experiment. The latter is used for determining significance of the changes.

increases at 1% yr^{-1} and then stabilizes at levels corresponding to 1 or 2 W m^{-2} of RF for 8 decades (see figure S11 for a depiction of GAT and RF associated with the two stabilization experiments). We focus on changes associated with ΔRF of 0.5 W m^{-2}, which in CESM1 produces statistically significant temperature change over 30% of the land surface when measured relative to the base period in a 1% CO_2 increase per year experiment, as in figure 3. We find that, depending on where along the two stabilization pathways this radiative forcing difference is measured, the fraction of land surface experiencing significant change can be as high as 85%. The fraction increases as the window within which the radiative forcing difference occurs moves along the forcing pathway. For example, it increases to 60% when the window spans the times at which radiative forcing first reaches 1.5 and 2.0 W m^{-2}. The fraction also increases as the window widens to include longer periods of stabilization at constant forcing. For example, it increases to 85% when the window begins at 1.5 W m^{-2} and includes the entire stabilization period at 2 W m^{-2}. The changes in GAT associated with these different measures of the same ΔRF of 0.5 W m^{-2} range from 0.31–0.63 °C. This

result is explained by inertia in the climate system which leads to a delay between reaching a given level of RF and its climate system consequences, so that stabilization of RF does not translate into an immediate stabilization of GAT (Collins *et al* 2013). We therefore emphasize that results based on differences in GAT are more reliable (less path dependent), as well as more precise, than those based on differences in RF.

Discussion and conclusions

The approach presented here could be applied to additional variables and/or tailored to particular regions of the world. We have used 20-year mean values of surface temperature and precipitation because of the common use of these measures in summarizing forced changes, for example in the projection chapters of IPCC reports (Collins *et al* 2013), as they represent a simple and agreed upon definition of climatological means. Other measures and other variables would have produced different quantitative results, but we propose our analysis of mean temperature and precipitation as a logical first step in characterizing regional changes implied by
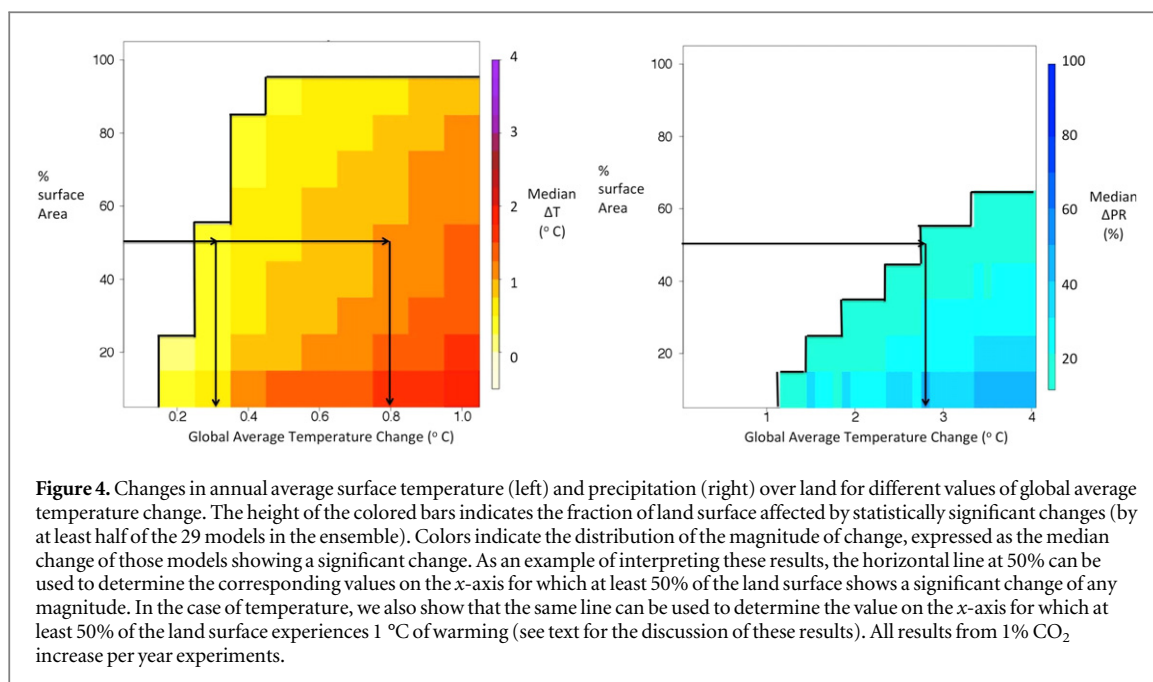
Figure 4. Changes in annual average surface temperature (left) and precipitation (right) over land for different values of global average temperature change. The height of the colored bars indicates the fraction of land surface affected by statistically significant changes (by at least half of the 29 models in the ensemble). Colors indicate the distribution of the magnitude of change, expressed as the median change of those models showing a significant change. As an example of interpreting these results, the horizontal line at 50% can be used to determine the corresponding values on the *x*-axis for which at least 50% of the land surface shows a significant change of any magnitude. In the case of temperature, we also show that the same line can be used to determine the value on the *x*-axis for which at least 50% of the land surface experiences 1 °C of warming (see text for the discussion of these results). All results from 1% CO_2 increase per year experiments.

different global measures. Our results suggest that regional outcomes in terms of average surface temperature start to present detectable differences—beyond internal variability—over a majority of the land surface, and for most models, starting at about 0.3–0.4 °C ΔGAT depending on whether annual or seasonal averages are considered. If substantial changes in regional temperature, rather than simple statistical significance, are of interest, for example warming upward of 1 °C locally, changes of at least 0.8 °C in global average temperature are required. Differences of these magnitudes are however not likely to produce significant changes in regional precipitation over most of the land surface. To achieve that type of change in annual average precipitation, a change in GAT of about 2.5 °C is required. Seasonal changes in precipitation require larger differences, of about 3.5 °C.

Since other types of variables, measured over different regional and time scales, would likely produce different quantitative results, we cannot establish an absolute criterion for the separation of scenarios when designing future climate model experiments. That design choice will have to reconcile demands from impact and mitigation communities with limited computational resources and time. Our conclusions regarding regional temperature differences, however, suggest that for some applications, climate model simulations based on forcing scenarios that lie between the existing representative concentration pathways would be worthwhile to pursue. It is also the case that emulator techniques (pre-eminent among which is pattern scaling; Santer *et al* 1990) are being proposed in order to interpolate climate model results between existing climate model simulations. In this study we do not assess the efficacy of these alternatives

to climate model simulations. It is possible that future developments of these types of empirical models will satisfy the needs of the impact and mitigation research community[3], but our results do not speak to the accuracy and promise of these methods, as we consider them outside the scope of our study.

Results also suggest that for measuring the consequences of exceeding a particular target, or anticipating the differences between alternative scenarios, differences in global average temperature (rather than radiative forcing) are the best predictor of differences in regional outcomes. This may appear obvious to part of the community interested in scenarios, but it is also obvious that many arguments, choices and debates are centered around differences in radiative forcing pathways and stabilization levels, as was the case in the definition of the RCPs (Moss *et al* 2008, 2010). Our results suggest that a better exploration of the potential implications of scenario differences could be carried out by first deriving the global average temperature differences implied by those different radiative forcing, which can be easily done by using simple climate models that do not require large computational efforts.

Acknowledgments

[3] A meeting to take stock of the current state of emulator techniques and synthesize the needs of the impact research community for climate information in order to prioritize future developments of these techniques was held at NCAR in April 2014. The meeting report can be downloaded from https://www2.image.ucar.edu/sites/default/files/event/PS2014WorkshopReport_0.pdf.

References

Baettig M B, Wild M and Imboden D M 2007 A climate change index: where climate change may be most prominent in the 21st century *Geophys. Res. Lett.* **34** L01705

Clarke L *et al* 2007 Scenarios of greenhouse gas emissions and Atmospheric concentrations. Sub-report 2.1A of synthesis and assessment product 2.1 by the US climate change science program and the subcommittee on global change research *Department of Energy, Office of Biological & Environmental Research* (Washington, DC, USA) pp 154

Clarke L *et al* 2014 Assessing transformation pathways *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* ed O Edenhofer *et al* (Cambridge: Cambridge University Press)

Collins M *et al* 2013 Long term climate change: projections commitments and irreversibility ed T F Stocker *Climate Change 2013: the Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge: Cambridge University Press)

Deser C *et al* 2012 Uncertainty in climate change projections: the role of internal variability *Clim. Dyn.* **38** 527–46

Frieler K *et al* 2013 Limiting global warming to 2 degrees C is unlikely to save most coral reefs *Nat. Clim. Change* **3** 165–70

Giorgi F and Bi X 2009 Time of emergence (TOE) of GHG-forced precipitation change hot-spots *Geophys. Res. Lett.* **36** L06709

Hawkins E and Sutton R 2009 The potential to narrow uncertainty in regional climate predictions *BAMS* **90** 1095–107

Hawkins E and Sutton R 2012 Time of emergence of climate signals *Geophys. Res. Lett.* **39** L01702

Held I and Soden B 2006 Robust responses of the hydrological cycle to global warming *J. of Climate* **19** 5686–99

Huntingford C *et al* 2012 The link between a global 2 degrees C warming threshold and emissions in years 2020, 2050 and beyond *Environ. Res. Lett.* **7** 014039

Hurrell J W *et al* 2013 The community Earth system model: a framework for collaborative research *B. Am. Meteorol. Soc.* **94** 1339–60

Lowe J A *et al* 2009 How difficult is it to recover from dangerous levels of global warming? *Environ. Res. Lett.* **4** 014012

Mahlstein I *et al* 2011 Early onset of significant local warming in low latitude countries *Environ. Res. Lett.* **6** 034009

Mahlstein I *et al* 2012 Perceptible changes in regional precipitation in a future climate *Geophys. Res. Lett.* **39** L05701

Manoj J *et al* 2011 Projections of when temperature change will exceed 2 degrees C above pre-industrial levels *Nat. Clim. Change* **1** 407–12

Meehl G A *et al* 2013 Climate change projections in CESM1(CAM5) compared to CCSM4 *J. Climate* **26** 6287–308

Meehl G A *et al* 2014 Climate model intercomparisons: preparing for the next phase *EOS Trans. Am. Geophys. Union* **95** 77–8

Meinshausen M *et al* 2009 Greenhouse-gas emission targets for limiting global warming to 2 degrees C *Nature* **458** 1158–U96

Moss R H *et al* 2008 *Towards New Scenarios for Analysis of Emissions, Climate Change, Impacts, and Response Strategies* IPCC Expert Meeting Report (Geneva: IPCC)

Moss R H *et al* 2010 The next generation of scenarios for climate change research and assessment *Nature* **463** 747–56

National Research Council 2011 *Climate Stabilization Targets: Emissions, Concentrations, and Impacts over Decades to Millennia* (Washington, DC: National Academies Press)

New M, Liverman D, Schroder H and Anderson K 2011 Four degrees and beyond: the potential for a global temperature increase of four degrees and its implications *Phil. Trans. of the Royal Soc.* **369** 6–19

Oppemheimer M *et al* 2014 Emergent risks and key vulnerabilities *Climate Change 2014: Impacts, Adaptation, and Vulnerability: A. Global and Sectoral Aspects. Contribution of Working Group 2 to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* ed C Field *et al* (Cambridge: Cambridge University Press) pp 1132

Rogelj J *et al* 2011 Emission pathways consistent with a 2 degrees C global temperature limit *Nat. Clim. Change* **1** 413–8

Rogelj J *et al* 2014 Probabilistic cost estimates for climate change mitigation *Nature* **506** 396–396

Santer B D, Wigley T M L, Schlesinger M E and Mitchell J F B 1990 Developing climate scenarios from equilibrium GCM results *Report No. 47* Max Planck Institute für Meteorologie, Hanburg, Germany (www.mpimet.mpg.de/fileadmin/publikationen/Reports/Report_47.pdf)

Schaeffer M *et al* 2012 Long-term sea-level rise implied by 1.5 degrees C and 2 degrees C warming levels *Nat. Clim. Change* **2** 867–70

Sedlacek J and Knutti R 2014 Half of the world population experience robust changes in water cycle for a 2 degrees warmer world *Environ. Res. Lett.* **9** 044008

Taylor K E, Stouffer R J and Meehl G A 2012 An overview of CMIP5 and the experiment design *BAMS* **93** 485

Tebaldi C and Friedlingstein P 2013 Delayed detection of climate mitigation benefits due to climate inertia and variability *PNAS* **110** 17229–34

UNFCCC 2009 Decision 2/CP.15 Copenhagen Accord *Report of the Conference of the Parties on its fifteenth session, Copenhagen, December 2009* (http://unfcc.int/resource/docs/2009/cop15/eng/11a01.pdf)

van Vuuren D P *et al* 2014 A new scenario framework for climate change research: scenario matrix architecture *Climatic Change* **122** 373–86

Waldhoff S *et al* 2014 Overview of the special issue: a multi-model framework to achieve consistent evaluation of climate change impacts in the United States *Climatic Change* (doi:10.1007/s10584-014-1206-0)

Wigley T M L 2006 A combined mitigation/geoengineering approach to climate stabilization *Science* **314** 452–4