

Models agree on forced response pattern of precipitation and temperature extremes

Article

Published Version

Fischer, E. M., Sedláček, J., Hawkins, E. ORCID:
<https://orcid.org/0000-0001-9477-3677> and Knutti, R. (2014)
Models agree on forced response pattern of precipitation and
temperature extremes. *Geophysical Research Letters*, 41 (23).
pp. 8554-8562. ISSN 0094-8276 doi:
<https://doi.org/10.1002/2014GL062018> Available at
<https://centaur.reading.ac.uk/40257/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1002/2014GL062018>

Publisher: American Geophysical Union

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online



RESEARCH LETTER

10.1002/2014GL062018

Key Points:

- Model agreement on changes in extremes is high in the absence of natural variability
- Pattern of forced response is robust for temperature and precipitation extremes
- Forced response is more robust for heavy precipitation than mean precipitation

Supporting Information:

- Readme
- Table S1 and Figures S1–S8

Correspondence to:

E. M. Fischer,
erich.fischer@env.ethz.ch

Citation:

Fischer, E. M., J. Sedláček, E. Hawkins, and R. Knutti (2014), Models agree on forced response pattern of precipitation and temperature extremes, *Geophys. Res. Lett.*, 41, 8554–8562, doi:10.1002/2014GL062018.

Received 25 SEP 2014

Accepted 17 NOV 2014

Accepted article online 20 NOV 2014

Published online 11 DEC 2014

Models agree on forced response pattern of precipitation and temperature extremes

E. M. Fischer¹, J. Sedláček¹, E. Hawkins², and R. Knutti¹
¹Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland, ²NCAS-Climate, Department of Meteorology, University of Reading, Reading, UK

Abstract Model projections of heavy precipitation and temperature extremes include large uncertainties. We demonstrate that the disagreement between individual simulations primarily arises from internal variability, whereas models agree remarkably well on the forced signal, the change in the absence of internal variability. Agreement is high on the spatial pattern of the forced heavy precipitation response showing an intensification over most land regions, in particular Eurasia and North America. The forced response of heavy precipitation is even more robust than that of annual mean precipitation. Likewise, models agree on the forced response pattern of hot extremes showing the greatest intensification over midlatitudinal land regions. Thus, confidence in the forced changes of temperature and precipitation extremes in response to a certain warming is high. Although in reality internal variability will be superimposed on that pattern, it is the forced response that determines the changes in temperature and precipitation extremes in a risk perspective.

1. Introduction

More frequent and intense climatic extremes are considered as a manifestation of climate change that would have severe socioeconomic and ecological impacts [Seneviratne *et al.*, 2011]. Models consistently project increases in global average quantities of heavy precipitation intensity and frequency along with rising temperatures [Collins *et al.*, 2013], which is consistent with physical considerations [Allan and Soden, 2008; Lenderink and Van Meijgaard, 2008]. For the global mean, the magnitude of the projected change is dependent on the model used, but there is strong agreement across the models over the direction of change [Tebaldi *et al.*, 2006; Min *et al.*, 2011; Sillmann *et al.*, 2013]. However, the geographical pattern of changes in heavy precipitation intensity by the middle or the end of the 21st century is highly uncertain, and at the grid point scale model simulations strongly differ in magnitude and sometimes even in sign [Fischer *et al.*, 2013; Kharin *et al.*, 2013; Sillmann *et al.*, 2013]. For a given emission scenario these uncertainties arise from parametric and structural model uncertainties and from internal variability.

For mean precipitation changes, disagreement in sign is often found where projected changes are small and still within the modeled range of internal variability, that is, where a response to anthropogenic forcings has not yet emerged locally in a statistically significant way [Schaller *et al.*, 2011; Tebaldi *et al.*, 2011; Power *et al.*, 2012]. As a consequence the area fraction of model agreement is found to increase only if the signal increases [Knutti and Sedláček, 2013].

For hot and cold temperature extremes there is a reasonable model agreement on the overall spatial pattern by the end of the 21st century [Orlowsky and Seneviratne, 2011; Sillmann *et al.*, 2013] and overall simulated trends have been found to be consistent with observed trends [Sillmann *et al.*, 2014]. However, high variability is superimposed upon the signal, which represents, for instance, a challenge to model evaluation of local changes in the observational period [Perkins and Fischer, 2013; Fischer and Knutti, 2014]. Likewise, heavy precipitation occurs irregularly and has a substantially higher year-to-year variability than mean precipitation. Thus, internal variability is the dominant uncertainty contribution for the intensification of precipitation extremes at local to regional scale even by the mid-21st century [Fischer *et al.*, 2013]. Only when aggregated over continental to global scale, the models agree on the fractional changes in the intensity of changes in extremes across the globe [Fischer *et al.*, 2013]. Put simply, models agree on the fraction of land that experiences a certain near-term intensification in heavy precipitation but disagree where they occur.

Does this imply that models show no robust response in the geographical patterns of heavy precipitation intensification to global warming? The answer depends on the definition of robustness or model agreement,

which is often ambiguous. For a given emission scenario two model simulations may disagree due to (1) a different underlying forced model response or (2) due to a different realization of internal variability [e.g., *Hawkins and Sutton*, 2009]. It is widely acknowledged that a single transient simulation of a coupled climate model should not be expected to follow the time evolution of the observations, due to the lack of predictability of internal variability at multidecadal timescales [*Branstator and Teng*, 2010; *Branstator et al.*, 2012; *Meehl et al.*, 2014]. Likewise, individual simulations for the future even from the same model do not agree in the presence of internal variability [*Deser et al.*, 2012a; *Fischer et al.*, 2013; *Deser et al.*, 2014]. Many multimodel intercomparison studies only take into account one realization per model and define robustness as agreement over the outcome of an individual realization, e.g., the change projected for the real world in 2080–2099. Alternatively, one can quantify agreement on the underlying long-term response in the absence of internal variability, the forced response. This ambiguity is not relevant for century-scale global mean temperature projections where agreement on individual realizations and agreement on forced response may be very similar, because the variability contribution is small. However, for regional intensification of heavy precipitation extremes robustness statements are distinctly different for the two definitions.

The forced response and a single realization also differ in terms of their spatial heterogeneity. At the regional scale even multimodel mean patterns of future heavy precipitation intensification look patchy [*Goubanova and Li*, 2007; *Radermacher and Tomassini*, 2012; *Rajczak et al.*, 2013; *Vautard et al.*, 2014]. This may either point to a high spatial heterogeneity in the physical processes leading to intensified heavy precipitation or result from a strong internal variability. Here we isolate the pattern of changes in heavy precipitation intensity and hot extremes in the absence of internal variability and assess for which areas and variables these patterns are robust across models.

2. Model Experiment and Observational Data

We analyze daily and monthly output of historical simulations for the period 1901–2005 as well as future projections forced with RCP8.5 for the period 2006–2100. We use output of 15 CMIP5 models that provide all the necessary output to analyze changes in hot and heavy precipitation extremes (see Table S1 in the supporting information). To avoid an overestimation of model agreement due to obvious model dependencies coming from the same modeling centers, we only use one model per modeling center by selecting the newest version or the one with the highest resolution. Thereby, we reduce but do not eliminate model dependencies [*Masson and Knutti*, 2011; *Knutti et al.*, 2013]. For the models HadGEM-ES, CSIRO-Mk-3-6-0, CanESM2, and EC-EARTH, for which four or more initial condition members are available (see Table S1), we use all the different realizations to test our estimates in a perfect model framework and to explore the benefits of multimember experiments to estimate the forced signal. The simulations are supplemented with a nine-member ensemble performed with the Community Earth System Model (CESM) version 1.0.4 including the Community Atmosphere Model version 4 (CAM4) and fully coupled ocean, sea ice, and land surface components [*Hurrell et al.*, 2013]. The nine members are initialized from different states of the preindustrial control simulation and differ in their initial conditions of the ocean, atmosphere, sea ice, and land components. Thereby, the nine CESM members use the exact same setup as the CMIP5 simulations.

The changes in mean and extremes of temperature and precipitation are expressed as local changes in degree C and percentage per degree C global warming. Agreement across models is quantified with two different metrics expressing (a) the agreement of the spatial pattern and (b) the agreement on the magnitude of the local change signals. The first metric is the area-weighted pattern correlation calculated for all possible pairs of models in the multimodel ensemble. The pattern correlation for all possible combinations of models is then averaged to obtain one estimate of pattern agreement across the experiments. The second metric quantifies the relative uncertainty against the mean signal, a ratio of uncertainty versus signal, similar to the signal-to-noise ratio. The relative uncertainty of the local signal is expressed as the ratio of the uncertainty, i.e., one standard deviation across the multimodel range and the multimodel mean change. First, this relative uncertainty metric is calculated at each grid point, and second, we calculate the area-weighted global median to get a robust global estimate. The local relative uncertainty metric may reach extremely high values when the local multimodel mean changes are nearly zero, but the global median is insensitive to these local outliers and yields a robust global metric of model agreement on the climate change signal.

Change in heavy precipitation intensity (Rx1day)

Historical (1986–2005 wrt 1901–20)

Forced signal (1901–2100)

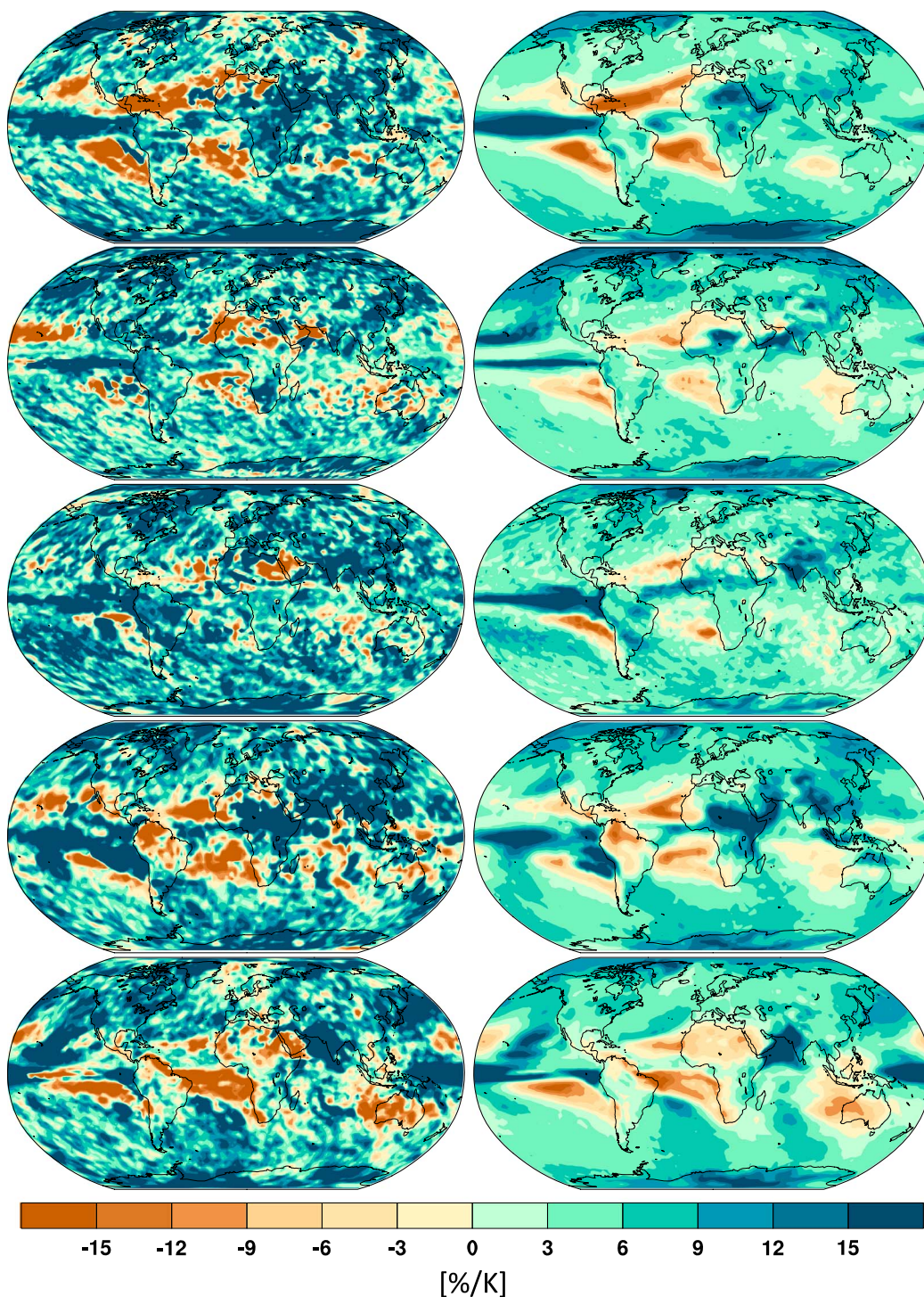


Figure 1. Model agreement on the change in heavy precipitation intensity in individual realizations and forced signal: (left) Change in 20 year means of annual 1 day precipitation maxima (*Rx1day*) in 1986–2005 with respect to 1901–1920 as simulated by the first member of CESM1-CAM4, HadGEM2-ES, EC-EARTH, CanESM2, and CSIRO-Mk3-6-0. Changes are expressed as local percentage changes per degree multimodel mean global warming. (right) Annual *Rx1day* per degree global warming of the respective model derived from a linear regression for the period 1901–2100. Regression slopes are averaged across 4–10 initial condition members of the same models.

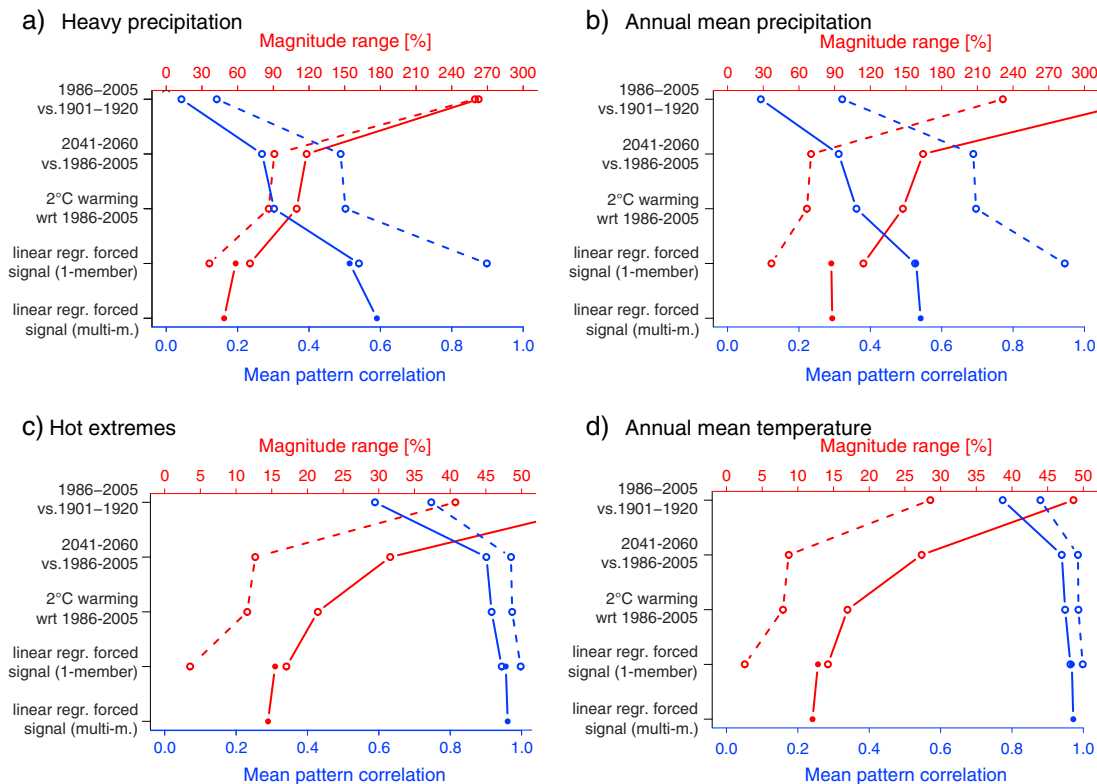


Figure 2. Model agreement for changes in mean and extremes of precipitation and temperature: (blue lines) Pattern agreement expressed as average of the pattern correlations (blue solid) across all combinations of 15 CMIP5 GCMs and (blue dashed) across members of the same model where available. (red lines) Agreement on the magnitude of the climate change signal expressed as the area-weighted median of the relative uncertainty range calculated at each grid point. The grid point scale uncertainty range is expressed as the one standard deviation across (red solid) multimodel range divided by the multimodel mean change and (red dashed) one standard deviation across the multimember range divided by the multimember mean change. The latter expresses the highest agreement possible given the level of noise in a perfect model case. Red and blue filled dots indicate the estimates for the five models shown in Figure 1 for which at least four members are available and open dots the estimates for 15 GCMs with one member available.

We focus on two extreme indices recommended by the World Meteorological Organization CCI/Climate Variability and Predictability/JCOMM Expert Team on Climate Change Detection and Indices [Zhang *et al.*, 2011], the intensity of hot extremes (TXx , the annual maximum of the daily maximum temperature), and heavy precipitation intensity ($Rx1day$, the maximum 1 day precipitation in a year).

3. Results and Discussion

Individual climate model simulations strongly disagree on the pattern of heavy precipitation changes during the twentieth century (1986–2005 relative to 1901–1920) shown in Figure 1 (left). Positive and negative changes show hardly any coherent pattern. Not a single grid point meets a rigorous model agreement criterion, defined as agreement on the sign of local changes in 80% of the 15 models (not shown). Does this imply that the simulations disagree because of fundamental differences in the way different models describe complex local processes and feedbacks controlling changes in heavy precipitation? We here argue that whether models agree or not depends on the definition of model agreement, which is often used in two ambiguous ways. Agreement on the intensification of heavy precipitation is poor for individual realizations of the twentieth century (Figure 1, left) but remarkably good on the forced signal of heavy precipitation intensity derived from multiple realizations of individual models (Figure 1, right). The forced signal of heavy precipitation is robust across models anywhere over land except for some arid to semiarid regions. In the following, we discuss the differences and the implications for the interpretation of model uncertainties. In order to systematically quantify agreement across CMIP5 models, we use two criteria, agreement on spatial pattern and agreement on the magnitude of changes (see Methods for details), and illustrate them for weak to strong climate change signals in Figure 2.

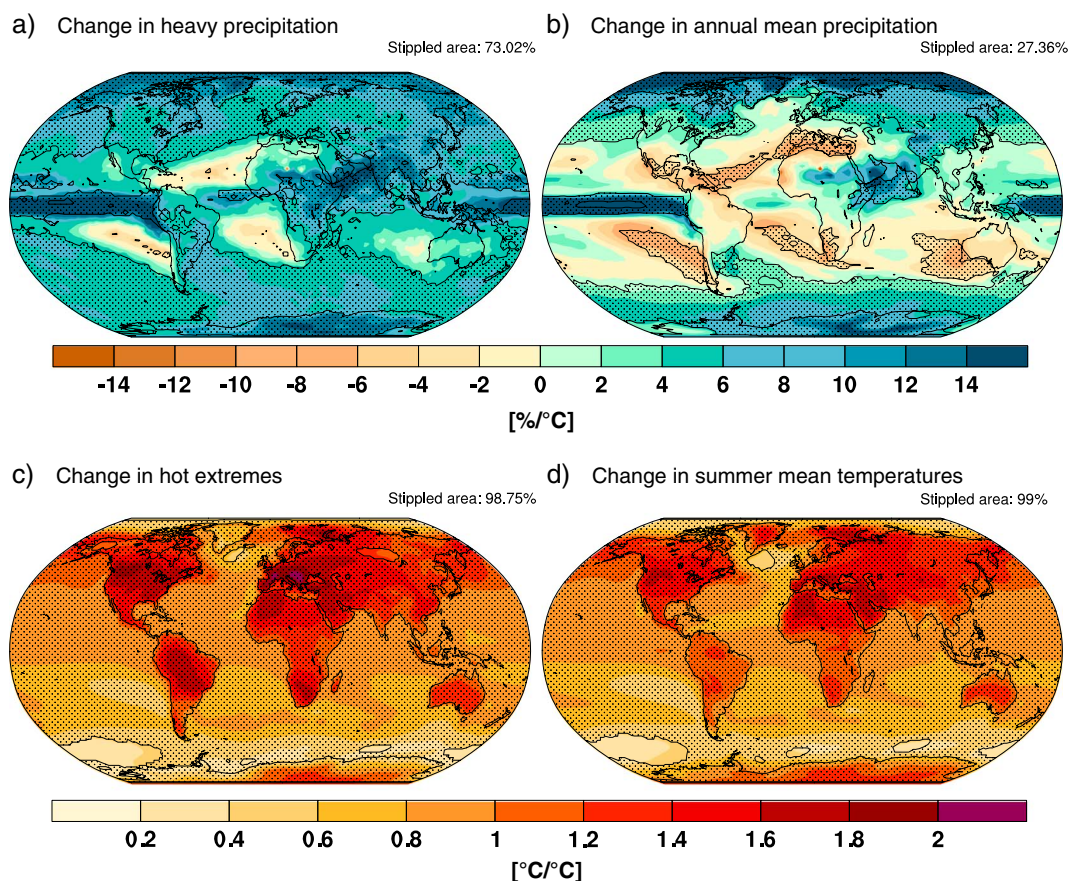


Figure 3. Model robustness in forced signal: Multimodel mean changes in (a) heavy precipitation intensity, (b) annual mean precipitation, (c) hot extremes, and (d) local summer mean temperature (June–July–August in Northern and December–January–February in Southern Hemisphere) per degree global warming in 15 CMIP models. Estimates are based on a linear regression of local changes with respect to global mean temperature change in the respective model simulation in the period 1901–2100 (historical and RCP8.5). Stippling illustrates agreement in sign of changes across at least 12 of the 15 models (80% of models).

Over the historical period (1986–2005 relative to 1901–1920) agreement on percentage changes in heavy precipitation per degree global warming is very poor. At most of the grid points the simulated changes even differ in sign so that the range across the 15 models is large at all grid points (Figure 2a, red solid line). This is not surprising since the disagreement at the grid point level across models and with observations of the past decades primarily results from large internal variability and lack of a strong signal [Fischer and Knutti, 2014].

The role of internal variability is less important once the climate change signal becomes more dominant, e.g., for the mid-21st century (2041–2060 with respect to 1986–2005). Changes are expressed as local percentage changes per multimodel mean warming (roughly 2°C in the multimodel mean). The mean pattern correlation across models is somewhat higher ($r = 0.30$) than for the historical period (Figure 2a, blue solid line) and the local spread is smaller (Figure 2a, red solid line and Figure S1) but the overall agreement is still poor. Internal variability is dominant for heavy precipitation changes by the midcentury, and even different realizations of the same model show a large spread (Figure 2a, dashed lines). One reason for the disagreement in the magnitude of heavy precipitation changes is that the models differ in their global warming by the mid-21st century. To account for the different rate of warming, we compare changes for the 20 year period in which each model reaches a global mean temperature increase of 2°C. This slightly improves the agreement in the magnitude of change (Figure 2a).

To understand to what extent the models agree in their response in the absence of any internal variability, one can extract the underlying forced signal by averaging multiple initial condition realizations of the exact

same model [Deser *et al.*, 2012a, 2012b, 2014]. However, such experiments are computationally costly and are only available for very few models. Here we use an alternative approach that maximizes the information from individual model simulations. To avoid the limitation to 20 year periods, we estimate the forced signal by a linear regression of the percentage change of annual 1 day heavy precipitation maxima (*Rx1day*) to the annual global mean temperature over the simulation period 1901–2100. The assumption is that the relative changes in heavy precipitation intensity scale linearly with the global mean temperatures (see discussion below). If the regression method is applied to several realizations of the same model, the patterns and magnitudes agree reasonably well (Figure 2 dashed blue and red lines), which confirms that this method efficiently filters much of the variability and yields a reasonable estimate of the forced signal.

The forced signal patterns estimated by linear regression agree remarkably well across the 15 models ($r = 0.51$, Figures 2a and S2). Likewise, the agreement on the magnitude of changes is much better (Figure S1) with a range that in many places is about 2–3 times smaller than expected from the 20 year differences at 2°C warming (Figure 2a). This indicates that for the CMIP5 models the patterns of forced heavy precipitation changes are remarkably consistent, with heavy precipitation becoming more intense over basically all land regions north of 35°N (Figure 3a). Likewise, models consistently simulate more intense precipitation across most of South America. For 73% of the land fraction at least 12 of the 15 models agree on the sign of the forced signal (Figure 3a). Over land, the discrepancies in the sign of the forced signal are largest over Australia, North Africa, and Central America. Those are a consequence of actual model differences in the forced response to global warming. Note that those are mostly arid to semiarid regions in which the wettest day per year is often not particularly extreme in an absolute sense.

Even when using the information for the 200 year period in the regression, the estimate of the forced signal is still affected by internal variability. This is demonstrated by the regression-based estimates of the forced signal from different simulations of the same model that agree well but not perfectly (Figure 2a, dashed lines). Thus, ideally, the regression-based estimate can be averaged across multiple realizations of the same model, in order to further average out the effect of internal variability. We therefore test the potential of averaging regression-based estimates across multiple realizations of the same model, using the five models shown in Figure 1 for which 4–10 realizations are available for the historical and RCP8.5 simulations from 1901 to 2100. The five models happen to agree on average somewhat less in their forced signal estimated from a single member than the larger set of 15 GCM (Figure 2a open versus closed circles). However, the agreement of the forced signal pattern derived from averaging regressions across several members (Figure 1, right) reveals an even higher model agreement ($r = 0.60$) than expected from single-member-based estimates, and the range across local changes in magnitude is comparatively small (Figure 2a, closed circles). The agreement is nearly as high for 5 day accumulated maximum precipitation (Figure S5). In summary, the model agreement increases, the stronger the signal becomes, the more information of the time series is included and the more simulations are averaged, an evolution that is summarized in Figure 2.

Annual mean precipitation has lower internal variability than heavy precipitation intensity and thus shows better agreement even if the signal is not very pronounced such as for historical changes or changes at 2°C global warming (Figure 2b). Nevertheless, even for annual mean precipitation, the robustness of the forced response is substantially underestimated based on 20 year periods at 2°C warming (Figure 2b). This is consistent with previous studies demonstrating that model disagreement on precipitation changes across multimodel experiments is hampered by the internal variability [Schaller *et al.*, 2011; Tebaldi *et al.*, 2011]. Only if the signal-to-noise ratio becomes large, model projections become more consistent [Knutti and Sedlacek, 2013]. Likewise, we here find that the regression-based estimates of the forced signal are in reasonable good agreement across models (Figure 2b). The similarity of the patterns becomes particularly evident if multiple members of individual models are considered (Figure S3) [Deser *et al.*, 2014].

Interestingly, we find that the pattern of the forced response is more consistent for heavy precipitation than mean precipitation (Figure 2a versus Figure 2b) despite the fact that annual mean precipitation experiences much smaller year-to-year variability than heavy precipitation. Likewise, there is a substantially larger area fraction at which the models agree on the forced signal of heavy precipitation (73%) than of annual mean precipitation (27%) (Figure 3). Even the relative uncertainty on the magnitude of heavy precipitation changes is substantially smaller than for mean precipitation changes (Figure 2). While heavy precipitation is strongly controlled thermodynamically by an increased saturation water vapor pressure

at higher temperature [Allan and Soden, 2008], mean precipitation is sensitive to both dynamic and thermodynamic changes and constrained by longwave radiative cooling [Allen and Ingram, 2002].

We find that for hot and cold temperature extremes, agreement in the forced model response is also substantially underestimated based on individual realizations of the twentieth century (Figures 2c and S5). Models agree on the regression-based forced signal estimates that the greatest intensification of hot extremes per degree warming occur over midlatitude land regions (Figures 3c and S4). Pattern correlations are around 0.95 (Figure 2c) and almost as large for hot extremes as for the annual mean temperatures (Figure 2d). Likewise, for cold extremes the forced response is robust across models (Figure S5) with greatest changes per degree global warming expected over high latitudes (Figures S6 and S7). The changes in hot and cold extremes show a similar pattern as the summer mean (Figure 3) and winter mean temperature changes (Figure S6), respectively, but regionally tend to substantially exceed the rate of mean changes.

In general, the linearly regressed estimates from different realizations of the same model agree well both in their pattern and magnitude (dashed lines in Figure 2), which underlines that the internal variability is efficiently filtered and linear regression on the global mean temperature is a reasonable rough estimate of the forced signal. The linearity assumptions for relative changes in heavy precipitation are reasonably justified as here shown for several regions in Figure S8. It does not hold everywhere and is only reasonable as long as global temperature changes are relatively small. Amplification of hot extremes, e.g., by land surface feedbacks [Seneviratne et al., 2006; Christensen et al., 2008; Fischer and Schär, 2010] are not problematic as long as the feedbacks are quasi-linear. However, there are obvious limits to the linearity of this amplification, e.g., for soil moisture feedbacks if the wilting point is reached and soils are completely dry [Fischer and Schär, 2010; Bellprat et al., 2013]. Finally, there are areas where the sign of the precipitation signal may reverse with increasing warming, e.g., in case of an equatorward shift of the Intertropical Convergence Zone precipitation [Hawkins et al., 2014]. Despite these limitations, we argue that the regression-based estimates of the forced signal are powerful, but ideally, it could be extracted from large multimember ensembles for high-emission scenarios. To this end multiple realizations are critical and daily and other high-frequency precipitation and surface air temperature from these simulations necessary to thoroughly assess the forced response in extremes.

4. Conclusions

Projections of how regional to local temperature and precipitation extremes will change by the middle or the end of the 21st century in the real world are very uncertain [Fischer et al., 2013; Kharin et al., 2013; Sillmann et al., 2013]. In contrast, we have demonstrated that models are surprisingly consistent in their forced response to a certain level of global warming, i.e., the changes in the absence of internal variability that will ultimately emerge in the very long run. Models consistently show an intensification of heavy precipitation across almost all land regions of Eurasia and North America. Interestingly, we find that both the pattern and the magnitude of forced response are more robust for heavy precipitation than mean precipitation. This is consistent with arguments that the signal to ratio, and thus the detectability of a signal, is higher for heavy precipitation than for mean precipitation [Hegerl et al., 2004; Fischer and Knutti, 2014], and consequently, in many places the heavy precipitation should emerge earlier from internal variability than annual mean precipitation.

Likewise, for hot extremes the forced pattern is consistent across models, with greatest intensification of hot extremes over the continental midlatitudes and warming of the hottest days that substantially exceed the global mean temperature change. The difference is that the forced signal is unaffected by variability, whereas in individual simulations of the twentieth and 21st century the change is dominated by internal variability. The model agreement becomes only evident if changes are averaged over large areas [Sillmann et al., 2013], aggregated in spatial probability density functions [Fischer et al., 2013] or if the noise is efficiently removed to isolate the forced signal. It is sometimes argued that the forced response is irrelevant since also in reality it will be superposed by high internal variability. However, in a risk perspective, it is the forced signal that determines the probabilistic changes in return levels or return periods. The forced signal determines the change in the expected probability of heavy precipitation or hot extremes, which ultimately underlies any risk assessments, whereas the individual realizations correspond to the actual outcome.

These results on one hand highlight the importance of large multimember ensembles for high-emission scenarios that allow for a robust isolation of the forced signal through averaging a large number of runs. Only this allows for an assessment of the robustness of projections and an isolation of actual model differences. On the other hand, our findings in a broader sense also show the importance of specifying whether model agreement or robustness refers to the forced signal or to individual realization. Likewise, the level of confidence in projections of extremes, often given in assessment reports, depends whether a statement applies to a single realization of the future or to the forced signal. Confidence in local changes of heavy precipitation projected for the midcentury anywhere in Eurasia or North America is relatively low, but this is not because models disagree on the response in heavy precipitation to a warming climate but rather because the signal-to-noise ratio is small and internal variability may obscure or even reverse the change. In contrast, confidence in the forced signal is high even at local to regional scales. Thus, it is very likely that in the long run heavy precipitation and hot extremes intensify anywhere in Eurasia or North America if the warming is large enough. Despite the agreement in the forced signal, models potentially share common deficiencies and need to be further scrutinized with observations and our theoretical understanding. Heavy precipitation is sensitive to the parameterization of convection, which involves large uncertainties. At cloud-resolving scales the response of hourly heavy precipitation intensity may be different to all the models used here [Kendon *et al.*, 2014]. Changes in temperature extremes are sensitive to changes in atmospheric blocking as well as land surface feedbacks, the representation of which is still deficient in current climate models. Nevertheless, our findings demonstrate that amid all the complexity of nonlinear processes controlling changes in hot and cold extremes and heavy precipitation, there may be simplicity as proposed by Held [2014]. Here a remarkably simple first-order pattern emerges: as global temperatures increase, the forced intensification of hot extremes and heavy precipitation is widespread over most of the land region and consistency across models for these changes is high, in particular over Eurasia and North America.

Acknowledgments

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. The data used in this study are available through the Coupled Model Intercomparison Project Phase 5 at <http://cmip-pcmdi.llnl.gov/cmip5/>.

The Editor thanks two anonymous reviewers for their assistance in evaluating this paper.

References

- Allan, R., and B. Soden (2008), Atmospheric warming and the amplification of precipitation extremes, *Science*, 321(5895), 1481–1484.
- Allen, M., and W. Ingram (2002), Constraints on future changes in climate and the hydrologic cycle, *Nature*, 419(6903), 224–232.
- Bellprat, O., S. Kotlarski, D. Lüthi, and C. Schär (2013), Physical constraints for temperature biases in climate models, *Geophys. Res. Lett.*, 40, 4042–4047, doi:10.1002/grl.50737.
- Branstator, G., and H. Teng (2010), Two limits of initial-value decadal predictability in a CGCM, *J. Clim.*, 23(23), 6292–6311.
- Branstator, G., H. Teng, G. Meehl, M. Kimoto, J. Knight, M. Latif, and A. Rosati (2012), Systematic estimates of initial-value decadal predictability for six AOGCMs, *J. Clim.*, 25(6), 1827–1846.
- Christensen, J. H., F. Boberg, O. B. Christensen, and P. Lucas-Picher (2008), On the need for bias correction of regional climate change projections of temperature and precipitation, *Geophys. Res. Lett.*, 35, L20709, doi:10.1029/2008GL035694.
- Collins, M., et al. (2013), Long-term climate change: Projections, commitments and irreversibility, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. F. Stocker et al., Cambridge Univ. Press, Cambridge, U. K., and New York.
- Deser, C., R. Knutti, S. Solomon, and A. Phillips (2012a), Communication of the role of natural variability in future North American climate, *Nat. Clim. Change*, 2(11), 775–779.
- Deser, C., A. Phillips, V. Bourdette, and H. Teng (2012b), Uncertainty in climate change projections: The role of internal variability, *Clim. Dyn.*, 38(3–4), 527–546.
- Deser, C., A. Phillips, M. Alexander, and B. Smoliak (2014), Projecting North American climate over the next 50 years: Uncertainty due to internal variability, *J. Clim.*, 27(6), 2271–2296.
- Fischer, E. M., and C. Schär (2010), Consistent geographical patterns of changes in high-impact European heatwaves, *Nat. Geosci.*, 3, 398–403.
- Fischer, E. M., and R. Knutti (2014), Detection of spatially aggregated changes in temperature and precipitation extremes, *Geophys. Res. Lett.*, 41, 547–554, doi:10.1002/2013GL058499.
- Fischer, E. M., U. Beyerle, and R. Knutti (2013), Robust spatially aggregated projections of climate extremes, *Nat. Clim. Change*, 3(12), 1033–1038.
- Goubanova, K., and L. Li (2007), Extremes in temperature and precipitation around the Mediterranean basin in an ensemble of future climate scenario simulations, *Global Planet. Change*, 57(1–2), 27–42.
- Hawkins, E., and R. Sutton (2009), The potential to narrow uncertainty in regional climate predictions, *Bull. Am. Meteorol. Soc.*, 90(8), 1095–1107.
- Hawkins, E., M. Joshi, and D. Frame (2014), Wetter then drier in some tropical areas, *Nat. Clim. Change*, 4(8), 646–647.
- Hegerl, G., F. Zwiers, P. Stott, and V. Kharin (2004), Detectability of anthropogenic changes in annual temperature and precipitation extremes, *J. Clim.*, 17(19), 3683–3700.
- Held, I. (2014), Simplicity amid complexity, *Science*, 343(6176), 1206–1207.
- Hurrell, J. W., M. Holland, P. Gent, S. Ghan, J. Kay, P. Kushner, J.-F. Lamarque, W. Large, D. Lawrence, and K. Lindsay (2013), The community Earth system model: A framework for collaborative research, *Bull. Am. Meteorol. Soc.*, 94, 1339–1360.
- Kendon, E. J., N. M. Roberts, H. J. Fowler, M. J. Roberts, S. C. Chan, and C. A. Senior (2014), Heavier summer downpours with climate change revealed by weather forecast resolution model, *Nat. Clim. Change*, 4, 570–576, doi:10.1038/nclimate2258.
- Kharin, V., F. Zwiers, X. Zhang, and M. Wehner (2013), Changes in temperature and precipitation extremes in the CMIP5 ensemble, *Clim. Change*, 11, 345–357.

- Knutti, R., and J. Sedláček (2013), Robustness and uncertainties in the new CMIP5 climate model projections, *Nat. Clim. Change*, 3(4), 369–373.
- Knutti, R., D. Masson, and A. Gettelman (2013), Climate model genealogy: Generation CMIP5 and how we got there, *Geophys. Res. Lett.*, 40, 1194–1199, doi:10.1002/grl.50256.
- Lenderink, G., and E. Van Meijgaard (2008), Increase in hourly precipitation extremes beyond expectations from temperature changes, *Nat. Geosci.*, 1(8), 511–514.
- Masson, D., and R. Knutti (2011), Climate model genealogy, *Geophys. Res. Lett.*, 38, L08703, doi:10.1029/2011GL046864.
- Meehl, G., et al. (2014), Decadal climate prediction: An update from the trenches, *Bull. Am. Meteorol. Soc.*, 95(2), 243–267.
- Min, S.-K., X. Zhang, F. W. Zwiers, and G. C. Hegerl (2011), Human contribution to more-intense precipitation extremes, *Nature*, 470(7334), 378–381.
- Orlowsky, B., and S. I. Seneviratne (2011), Global changes in extreme events: Regional and seasonal dimension, *Clim. Change*, 110, 669–696.
- Perkins, S. E., and E. M. Fischer (2013), The usefulness of different realizations for the model evaluation of regional trends in heat waves, *Geophys. Res. Lett.*, 40, 5793–5797, doi:10.1002/2013GL057833.
- Power, S., F. Delage, R. Colman, and A. Moise (2012), Consensus on twenty-first-century rainfall projections in climate models more widespread than previously thought, *J. Clim.*, 25(11), 3792–3809.
- Rademacher, C., and L. Tomassini (2012), Thermodynamic causes for future trends in heavy precipitation over Europe based on an ensemble of regional climate model simulations, *J. Clim.*, 25(21), 7669–7689.
- Rajczak, J., P. Pall, and C. Schär (2013), Projections of extreme precipitation events in regional climate simulations for Europe and the Alpine Region, *J. Geophys. Res. Atmos.*, 118, 3610–3626, doi:10.1002/jgrd.50297.
- Schaller, N., I. Mahlstein, J. Cermak, and R. Knutti (2011), Analyzing precipitation projections: A comparison of different approaches to climate model evaluation, *J. Geophys. Res.*, 116, D10118, doi:10.1029/2010JD014963.
- Seneviratne, S. I., D. Lüthi, M. Litschi, and C. Schär (2006), Land-atmosphere coupling and climate change in Europe, *Nature*, 443, 205–209.
- Seneviratne, S. I., et al. (2011), Changes in climate extremes and their impacts on the natural physical environment, in *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation, A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC)*, edited by C. B. Field et al., pp. 109–230, Cambridge Univ. Press, Cambridge, U. K., and New York.
- Sillmann, J., V. Kharin, F. Zwiers, X. Zhang, and D. Bronaugh (2013), Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections, *J. Geophys. Res. Atmos.*, 118, 2473–2493, doi:10.1002/jgrd.50188.
- Sillmann, J., M. G. Donat, J. C. Fyfe, and F. W. Zwiers (2014), Observed and simulated temperature extremes during the recent warming hiatus, *Environ. Res. Lett.*, 9(6), 064023, doi:10.1088/1748-9326/9/6/064023.
- Tebaldi, C., K. Hayhoe, J. M. Arblaster, and G. A. Meehl (2006), Going to the extremes, *Clim. Change*, 79(3–4), 185–211.
- Tebaldi, C., J. Arblaster, and R. Knutti (2011), Mapping model agreement on future climate projections, *Geophys. Res. Lett.*, 38, L23701, doi:10.1029/2011GL049863.
- Vautard, R., et al. (2014), The European climate under a 2 degrees C global warming, *Environ. Res. Lett.*, 9(3), doi:10.1088/1748-9326/9/3/034006.
- Zhang, X., L. Alexander, G. Hegerl, P. Jones, A. Tank, T. Peterson, B. Trewin, and F. Zwiers (2011), Indices for monitoring changes in extremes based on daily temperature and precipitation data, *Wiley Interdiscip. Rev. Clim. Change*, 2(6), 851–870.