Detection of spatially aggregated changes in temperature and precipitation extremes

E. M. Fischer and R. Knutti

1Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

Abstract

Observed trends in the intensity of hot and cold extremes as well as in dry spell length and heavy precipitation intensity are often not significant at local scales. However, using a spatially aggregated perspective, we demonstrate that the probability distribution of observed local trends across the globe for the period 1960–2010 is clearly different to what would be expected from internal variability. We detect a distinct intensification of heavy precipitation events and hot extremes. We show that CMIP5 models generally capture the observed shift in the trend distribution but tend to underestimate the intensification of heavy precipitation and cold extremes and overestimate the intensification in hot extremes. Using an initial condition experiment sampling internal variability, we demonstrate that much of the local to regional differences in trends of extremes can be explained by internal variability, which can regionally mask or amplify the forced long-term trends for many decades.

1. Introduction

Given the high societal, economic, and ecological impacts of recent climatic extremes, it is vital to understand how their intensity and frequency change with increasing greenhouse-gas concentrations. Robust evidence for observed changes in extremes is mostly based on globally aggregated metrics. For instance there is robust observational evidence that the global average intensity of hot and cold extremes and the number of heatwaves have significantly changed over the last decades [Alexander et al., 2006; Donat et al., 2013a; Perkins et al., 2012]. This is consistent with the observed shift in the global distribution of daily minimum and maximum temperatures [Donat and Alexander, 2012] and the fact that extremely warm months have changed with a greater amplitude than annual mean temperatures [Hansen et al., 2012]. Finally, aggregated globally and continentally the recent numbers of record-breaking events both at monthly and daily time scales are higher than expected in an unforced climate state [Coumou et al., 2013; Meehl et al., 2009; Rahmstorf and Coumou, 2011].

However, local to regional changes in extremes are more difficult to detect due to the high variability at small spatial scales, and evidence for significant changes is mostly based on exceedance numbers of moderate thresholds, such as a widespread decrease in the number of cold days and nights and an increase in warm nights and hot days [Donat et al., 2013a]. Likewise, evidence for regional changes in precipitation extremes is often based on moderate indices such as increasing number of days with precipitation higher than 10 mm or the simple daily intensity index [Alexander et al., 2006; Donat et al., 2013a].

However, in contrast to these moderately defined indices, trends since the mid-20th century in the intensity of hot and cold extremes, expressed as annual temperature maxima (TXx) or minima (TNn), are significant over far less than half of the regions with data availability [Alexander et al., 2006; Donat et al., 2013a]. Significant changes in heavy precipitation intensity, expressed as maximum 5 day accumulated rainfall (RX5day), have only been detected in relatively few regions [Donat et al., 2013a]. Finally, it has been pointed out that observed trends in severity of droughts are sensitive to the index definition, and using an updated methodology little change in drought has been observed over the past 60 years [Sheffield et al., 2012].

Despite the small fraction of changes, observed patterns of changes in temperature extremes have been detected and attributed to anthropogenic influence [e.g., Christidis et al., 2011; Christidis et al., 2005; Min et al., 2013; Morak et al., 2011]. Likewise, it has been shown that anthropogenic forcing has contributed to a global-scale intensification of heavy precipitation over the second half of the 20th century [Min et al., 2011; Zhang et al., 2013]. These detection and attribution studies are based on sophisticated optimal fingerprinting methodologies.
Here we propose a simple and intuitive alternative detection approach, which allows detecting whether observed trends differ from internal variability and evaluating whether the simulated and observed trends are consistent. In contrast to the optimal fingerprinting methods, our approach is limited to detection and is not designed to attribute the changes to specific forcings. Here we use our novel approach to explore to what extent the spatial distribution of trends in the intensity of extremes differs from the ones expected from internal variability and whether the observed changes are captured by climate models.

2. Model Experiment and Observational Data

We analyze historical simulations as well as the first years of the RCP8.5 simulations of 25 CMIP5 models (see Table S1). The simulations are compared with a multimember ensemble (described in detail in Fischer et al. [2013]) 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]. All simulations in the CESM multimember ensemble are driven with historical forcing until 2005 and with the RCP8.5 thereafter. On 1 January 1950, a small random perturbation in the order of $10^{-13}$ is imposed to the atmospheric initial condition field of the reference run to produce a 21-member initial condition ensemble covering the period 1950–2100. Fischer et al. [2013] show that about a decade after the initial perturbation the spread across the members is comparable to an experiment in which the members are initialized from different starting points in a preindustrial control simulation (i.e., different atmosphere and ocean initial conditions). The simulated extreme indices are evaluated against the GHCNDEX [Donat et al., 2013b] as well as the HadEX2 data set [Donat et al., 2013a] for the period 1960–2010. Annual mean trends are compared for temperatures with the high-resolution gridded data set CRUTS3.2 [Harris et al., 2013] and for precipitation with the GPCC Full Data Product version 6 with a 1 $\times$ 1° resolution [Rudolf et al., 2005]. For comparison against observational data the extreme indices in the models were calculated on their native grids and then interpolated to the grid of the GHCNDEX data set and masked to the grid points for which data were available for at least 49 years in the period 1960–2010. The analysis is restricted to land regions between 66°S and 66°N.

3. Results

The two gridded data sets, HadEX2 and GHCNDEX, consistently show an intensification of hot extremes (expressed as annual temperature maxima, TXX, see Supplementary Information) over large parts of Europe, parts of northern Asia (around Mongolia) by 2–3 K in the period 1960–2010 (Figures 1a and 1b). Over Australia and the Americas, only minor regions experienced a significant intensification of hot extremes. In contrast to the increase in warm nights (TN90p) and hot days (TX90p), which exhibit significant trends over most of the land fraction [Alexander et al., 2006; Donat et al., 2013a], the intensification of hot extremes (TXX) is only significant (tested with nonparametric Mann-Kendall test at 95% level) at about a third of the grid points where data are available (35% in HadEX2 and 32% in GHCNDEX, respectively). Likewise, the intensification of heavy precipitation events, here expressed as annual maxima of 5 day accumulated precipitation (RX5day), is only significant at few grid points (Figures 1e and 1f, about 10% in HadEX2 and 8% in GHCNDEX). The relatively small area of significance is not surprising given the very high internal variability at the grid point scale of these intensity measures and the comparatively short period.

However, if the trends are aggregated in a spatial probability perspective similar to the one proposed in Fischer et al. [2013] (see Supplementary Information), it is evident that there is a distinct shift in trend distribution. To illustrate this, the extreme indices at each grid point have been normalized by their interannual variability in the period 1986–2010. Grid point-scale trends are then globally aggregated in a spatial probability density function (PDF), which quantifies the land fraction exhibiting a certain trend. Local trends are weighted by their latitude-dependent area.

We find that the land fraction exhibiting positive observed trends in hot extremes (TXX) is much larger than the one showing negative trends (Figure 2a). More than a third of the land fraction experienced trends that are larger than 1σ over the period 1960–2010, or 0.2σ per decade. A crucial question is whether the observed trend PDF differs from what would be expected from internal variability. Even in an unforced climate multidecadal linear trends would differ from zero at most grid points. This is illustrated by the grey PDF, derived from multicentury control simulations from four climate models (CESM, HadGEM2-ES, MPI-ESM-MR,
Figure 1. Observed and simulated trends in intensity of hot and precipitation extremes. Linear trends in the intensity of (a–d) hot extremes (TXx) and (e–h) heavy precipitation (RX5day) in the period 1960–2010. Linear trends are calculated based on HadEX2 and GHCNDEX gridded data as well as member 7 and member 16 of the CESM-IC initial condition member ensemble. Stippling indicates significant trends at the 95% confidence level tested with a nonparametric Mann-Kendall test.
Fifty-one-year linear trends have been calculated from a large sample of time windows for each of the control simulations. The different control simulations yield a similar normalized trend distribution (Figure S1). Thus, we here combine all four control simulations to quantify the trend distribution expected from internal variability (5–95% range shown in light grey). The mean PDF of the control simulations is consistent with an alternative estimate based on randomly resampling the GHCNDEX observational data at each individual grid point and constructing new PDFs (dark grey).

**Figure 2.** Spatial distribution of trends in hot and cold extremes. Probability density function of the land fraction (66°S–66°N) experiencing a certain linear trend (1960–2010) in (a) hot and (b) cold extremes as well as annual mean temperatures. The green and yellow lines show the distribution of observed trends, and the red and blue band illustrate the 5–95th range across CMIP5 and CESM-IC, respectively, for each bin. The light and dark grey range illustrate the range expected due to internal variability estimated from a combination of four control simulations and resampling the observational data. The lines and grey band at the bottom mark the 5–95% range for the mean across each data set. Models are regridded to and masked with the availability of GHCNDEX data for extreme indices. See Figure S2 for the corresponding simulated trend distributions at all grid points. Trends have been normalized by the local interannual standard deviation 1986–2005.
We show that the land fraction exhibiting positive trends in hot extremes is clearly larger than expected from internal variability (Figure 2a). Likewise, the observed PDF of trends in cold extremes (TNn) clearly differs from what would be expected due to internal variability (Figure 2b). A substantial land fraction exhibits positive trends in cold extremes that are larger than 0.4σ/decade. Such trends are substantially larger than found in any period of the control runs or in resampled observational data series. The warming trends in hot and cold extremes are often larger in absolute terms than for annual mean temperatures. Nevertheless, the PDFs of normalized trends show a less distinct shift for the extremes than for annual mean temperatures (Figure 2c), primarily due to their higher year-to-year variability.

We also detect a distinct intensification of heavy precipitation events that emerges from internal variability (Figure 3a). While only few grid points exhibit a significant trend to more intense heavy precipitation (a), the PDF of trends in dry spell length (CDD) shows a clear shift (Figure 3b). Annual mean precipitation trends are less distinct (Figure 3c), primarily due to the higher year-to-year variability.

Figure 3. Spatial distribution of trends in heavy precipitation, dry spell length, and annual mean precipitation. Same as Figure 2 but for (a) heavy precipitation intensity, RX5day, (b) dry spell length, CDD, and (c) annual mean precipitation. Models are regridded to and masked with the availability of GHCNDEX data for extreme indices. See Figure S3 for the corresponding simulated trend distributions at all grid points.
(Figures 1e and 1f), the spatial distribution of observed trends (yellow and green) differs from the ones in preindustrial control simulations (light grey). The fractions with positive trends are clearly larger than with negative trends, which is consistent with the finding that a majority, about two third, of the observed station series exhibit positive trends in daily precipitation maxima [Westra et al., 2013]. Furthermore, Figure 3a reveals that particularly the area fraction, which has seen strongly positive trends (>0.2σ/decade), is substantially larger than expected from internal variability in control simulations. The detection is insensitive to the normalization of precipitation extremes and we also detect a signal if trends are expressed as percentage changes without normalization. Note that the area with data availability only covers a limited portion of the land fraction with high coverage over the northern midlatitude to high latitudes, where mean precipitation increases are expected to be most pronounced and robust [Knutti and Sedlacek, 2013]. However, in CMIP5 models this area is found to be representative for the whole globe and the intensification of heavy precipitation events over the entire land area between 66°N and 66°S is even more pronounced than for the area for which observational data are available (Figure S3). Our findings are based on a novel approach but consistent with early studies demonstrating a detection of changes in heavy precipitation with optimal fingerprinting methods [Min et al., 2011; Zhang et al., 2013].

Note that the emergence of observed trends in precipitation extremes is more pronounced than for annual mean precipitation (Figure 3c). Some studies have suggested that the response of precipitation extremes to warming temperatures is stronger than for mean precipitation [e.g., Allen and Ingram, 2002; Allan and Soden, 2008; Lenderink and Van Meijgaard, 2008; Kharin et al., 2013]. However, the internal variability (noise) is also substantially stronger for heavy precipitation. Consistent with the argument by Hegerl et al. [2004], we find that the signal-to-noise ratio and thus the detection of heavy precipitation intensity are clearer than for mean precipitation. The multimember CESM-IC mean forced signal in extreme precipitation is an intensification at almost any point, which manifests itself in shift of the PDF to higher precipitation extremes (Figure 3a), whereas mean precipitation exhibits a widening of the spatial PDF with both positive and negative trends that are larger than expected from internal variability (Figure 3c). This implies that we detect both a reduction and increase in mean precipitation. The positive trends in annual mean precipitation are clearly detected particularly over the generally wet northern midlatitudes to high latitudes (here defined as 45°N–66°N, Figure S4a), consistent with Zhang et al. [2007] and Polson et al. [2013]. The drying is detected primarily over the zonal bands between 15°S and 15°N (Figure S4b).

The climate change signal is least clear for dry spell length, here expressed as the annual maximum number of consecutive dry days (CDD). The observed distribution of trends is at the low end of the range explained by internal variability (Figure 3b). The observations indicate a weak tendency toward shorter dry spell length. However, in contrast to the intensification of heavy precipitation, models suggest that in this case the area with observational data availability is not representative for the total land area, for which the models simulate no discernible change (Figure S3b).

Since climate models are used for attribution studies and projections of extreme events, it is vital to understand whether they capture the observed trends in temperature and precipitation extremes. To evaluate this, we first test whether the observed trend patterns are reproduced by the CMIP5 simulations. Figure S5 reveals that some GCMs capture the observed regional intensification of hot extremes (TXx) in Asia and others the intensification trend over Europe. However, GCMs do not consistently reproduce the overall observed global trend pattern. Particularly for heavy precipitation, many of the GCMs simulate the changes over completely different regions than the observations.

Does this imply that GCMs suffer from systematic deficiencies in simulating the observed regional trends? Such an interpretation would be valid if the regional trend patterns would mostly result from external forcing. In that case all realizations of one model would show the same regional trend patterns. However, in the 21-member initial condition experiment CESM-IC we find very different regional trend patterns in the exact same model, forced with identical historical radiative forcing but slightly different atmospheric initial conditions. This is illustrated by a comparison of CESM-IC member 7 (Figure 1c), which partly captures the trends in hot extremes over Asia but simulates no significant trends in Europe, and member 16 (Figure 1d) that simulates no trends over eastern Asia but captures the warming trends in Europe. A comparison across different realizations of the same model shows a very large spatial heterogeneity in the trend patterns (Figure S6). The differences between members are remarkable given that forced signal of the model (approximated
as multimember mean trend) exhibits a significant intensification of hot extremes everywhere over land (Figure S7a). The forced signal also shows distinct regional differences with a maximum intensification over Europe and the eastern US. However, at the time scales of five decades, the forced trends are completely masked or amplified by internal variability. This is even more pronounced for heavy precipitation where trends vary strongly between realizations (Figure S8). In summary, this implies that at time scales of a few decades, the location of significant trends of hot extremes and particularly of heavy precipitation is largely determined by internal variability even if averaged across large regions (Figure S9), which is consistent with recent studies emphasizing the role of internal variability for observed and projected trends [Deser et al., 2012; Fischer et al., 2013; Perkins and Fischer, 2013].

Despite the large differences in regional trend patterns, models agree on the land fraction exhibiting certain trends. This becomes evident in the probability perspective shown in Figures 2 and 3 where CMIP5 models and CESM-IC members have been masked by the GHCNDEx grid points with available data. We find that the CMIP5 models (5–95th percentile range, light red) as well as the CESM-IC multimember ensemble (blue) show a consistent fraction exhibiting certain trend magnitudes in hot extremes (Figure 2a). Compared to observations many models tend to overestimate the observed trends in hot extremes (TXx), primarily because they show not enough grid points with negative trends. The changes in cold extremes (TNn) are found to be systematically and strongly underestimated by the CMIP5 models and CESM-IC members (Figure 2b), which is consistent with Min et al. [2013]. Likewise, the intensification of heavy precipitation events is underestimated by the models (Figure 3a). Finally, models underestimate the drying trend in annual mean precipitation, i.e., they show a too small portion of the globe with negative trends (Figure 3c).

The trend PDFs are narrower for simulations than for observations of annual means as well as of hot and cold extremes. This implies that the models show less spatial diversity in trends than observed, consistent with earlier results for annual and seasonal mean trends [Bhend and Whetton, 2013; van Oldenborgh et al., 2013].

The fact that the models capture the overall tendency of an intensification of hot and heavy precipitation extremes should also enhance our confidence in the underlying forced signal, suggesting an overall increase at most gridboxes that is expected to emerge in the long term [Tebaldi et al., 2006]. Note that our simple approach does not represent a comprehensive model trend evaluation, which is beyond the scope of this study and has been done before for climatological values of extremes [Sillmann et al., 2013].

### 4. Conclusions

Precipitation extremes have significantly intensified in the period 1960–2010 over roughly 10% and hot extremes over about 30% of the land fraction. Due to the high internal variability at the local scale and the limited period of five decades the trends are nonsignificant at most grid points. However, a spatially aggregated perspective reveals that the intensification of hot and heavy precipitation extremes (TXx and RX5day) as well as the warming of cold extremes (TNn) is substantially larger than expected from internal variability. These findings that are based on a novel and simple approach are consistent with recent studies that detected a significant change in temperature extremes [Morak et al., 2011; Christidis et al., 2005; Morak et al., 2013] as well as an in heavy precipitation intensity [Min et al., 2011; Zhang et al., 2013] based on optimal fingerprinting methods. We show that for hot and cold extremes as well as for heavy precipitation intensity, the land fraction exhibiting positive trends is clearly larger than for negative trends. Likewise, there is a larger fraction that has experienced a strong intensification of hot extremes and heavy precipitation of more than 0.20/decade than in any comparable period in control simulations.

We further demonstrate that we should not expect models to reproduce the spatial pattern of intensification in temperature and precipitation extremes as internal variability can locally obscure or amplify local trends for many decades, as discussed in Perkins and Fischer [2013] on the example of heatwave trends in Australia. Here we show that different realizations of the same model yield very different spatial trend patterns in extremes. However, the underlying mean signal across all members of the same model, that is expected to emerge in the long term, shows an intensification of hot and precipitation extremes and warming of cold extremes at almost all grid points. We show that the simulated shift in trend distributions are in agreement with observations, however, with some biases in the magnitude. Overall most models are found to underestimate the observed trends in cold extremes and heavy precipitation and overestimate the trends in hot extremes, which have been shown for CMIP3 [Christidis et al., 2011; Min et al., 2011; Zwiers et al., 2011].
In summary, despite the fact that trends in extremes are not significant at the majority of grid points, a spatially aggregated perspective yields a clear detectable signal. For precipitation extremes the observed shift in the trend distribution is more distinct than for annual mean precipitation.

Acknowledgments
We would like to thank Markus Donat, Lisa Alexander, and the whole CLIMDEX team for making available the gridded observational data sets at www.climdex.org through the WMO ETCCDI, Linkage Project LP100200690. We acknowledge the World Climate Research Programme’s Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modelling groups for producing and making available their model output. The Editor thanks two anonymous reviewers for assistance evaluating this manuscript.

References
Lenderink, G., and E. Van Meijgaard (2008), Increase in hourly precipitation extremes beyond expectations from temperature changes, Nat. Geosci., 1(8), 511–514.