Projections of global changes in precipitation extremes from CMIP5 models

Andrea Toreti,[1] Philippe Naveau,[2] Matteo Zampieri,[3] Anne Schindler,[1]

Enrico Scoccimarro,[3,4] Elena Xoplaki,[1] Henk A. Dijkstra,[5] Silvio Gualdi[3,4],

and Jürg Luterbacher[1]

Corresponding author: A. Toreti, Dept of Geography, Climatology, Climate Dynamics and Climate Change, Justus-Liebig University of Giessen, 35390 Giessen, Germany. (andrea.toreti@geogr.uni-giessen.de)

[1]Dept. of Geography, Climatology,

Precipitation extremes are expected to increase in a warming climate, thus it is essential to characterise their potential future changes. Here we evaluate eight high-resolution Global Climate Model simulations in the twentieth century and provide new evidence on projected global precipitation extremes for the 21st century. A significant intensification of daily extremes for all seasons is projected for the mid and high latitudes of both hemispheres at the end of the present century. This signal supports a dynamical influence of the polar warming amplification on precipitation extremes.For the subtropics and tropics, the lack of reliable and consistent estimations found for

Climate Dynamics and Climate Change,
Justus-Liebig University of Giessen,
Giessen, Germany.

[2]Laboratoire des Sciences du Climat et de l'Environnement, IPSL-CNRS,
Gif-sur-Yvette, France.

[3]Centro Euro-Mediterraneo sui Cambiamenti Climatici, Lecce, Italy

[4]Istituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy.

[5]Dept. of Physics and Astronomy, Utrecht University, Utrecht, The Netherlands.

both the historical and future simulations might be connected with model deficiencies in the representation of organised convective systems. Low inter-model variability and good agreement with high-resolution regional observations are found for the twentieth century winter over the Northern Hemisphere mid and high latitudes.

1. Introduction

Exposure and vulnerability to weather and climate-related natural hazards largely determine the severity of impacts of these extremes [*IPCC*, 2012]. In the context of climate change, where considerable changes in the frequency and intensity of extremes are expected, the development of adequate risk-reduction strategies and measures is crucial. Since planning requires reliable knowledge of the relevant climate phenomena, a robust characterisation in terms of frequency and intensity of current and future extreme precipitation is of great relevance. Changes in mean annual precipitation have been observed in different regions of the world, with decreasing tendencies over the tropics and subtropics of the Northern Hemisphere (NH), increases over the northern mid and high latitudes and over the tropics and subtropics of the Southern Hemisphere (SH) [*Trenberth et al.*, 2007; *Zhang et al.*, 2007; *Trenberth*, 2011]. A widespread increase both in the frequency and intensity of daily precipitation extremes has already been identified [*Alexander et al.*, 2006; *Min et al.*, 2011; *Westra et al.*, 2013], although uncertainties arising from the lack of observations at the daily scale and the interpolation procedures applied for the production of available gridded data sets affect the estimations of precipitation extremes [*Trenberth et al.*, 2007; *Chen and Knutson*, 2008; *Hofstra et al.*, 2009; *O'Gorman and Schneider*, 2009; *Min et al.*, 2011; *Trenberth*, 2011]. In the NH, this upward tendency (that has been identified in the second half of the twentieth century) has been linked to human-induced greenhouse gases increase [*Min et al.*, 2011]. Global Climate Models (GCMs) still cannot adequately capture the frequency, the intensity, the tendency and the spatial distribution of observed precipitation extremes over large regions in the world [*Sun et al.*, 2006; *Allan*

and Soden, 2008; *O'Gorman and Schneider*, 2009; *Min et al.*, 2011].

Global warming implies an increase of atmospheric water vapour content at a rate of about $7\%/K$, through the Clausius-Clapeyron equation [*Allan and Soden*, 2008]. Hence, a comparable increase in extreme precipitation would be expected over the next decades [*O'Gorman and Schneider*, 2009; *Kharin et al.*, 2013]. A significant reduction in return times of annual extremes of daily precipitation (20-year return level) has been globally projected for different radiative forcing scenarios for the late 21^{st} century with large inter-model disagreement in the tropics [*Kharin et al.*, 2013].

Here, we evaluate the simulated daily precipitation extremes in the 20^{th} century assuming stationary processes [e.g., *Scoccimarro et al.*, 2013]. This implies that the ability of the models to reproduce the observed tendencies in specific regions of the world is not considered. Furthermore, we provide for the first time a comprehensive global assessment of seasonal future changes in daily precipitation extremes identifying regions where both consistency (i.e., models agreement) and reliability (i.e., goodness-of-fit of the applied statistical model) are achieved.

2. Data and Methods

Simulations for the period 1966-2099 were retrieved from the Coupled Model Intercomparison Project CMIP5 [*Taylor et al.*, 2012]. Eight models with a horizontal atmospheric resolution higher than 1.5° were chosen (Table S1) and daily precipitation data were retrieved. As for the projections (2006-2099), the high emissions scenario RCP8.5 and the mid-range mitigation emissions scenario RCP4.5 [*Moss et al.*, 2010] were selected. With respect to observations, an equivalent global high-resolution daily precipitation data set

covering the historical period 1966-2005 does not exist. Therefore, freely available high resolution regional products were retrieved. As shown in the Table S2, these gridded products have a spatial resolution of 0.25° and 0.05° over Australia. They do not cover the entire world, but they provide a very good coverage of the Euro-Mediterranean region, Northern Eurasia, the Middle East, Asia, Australia and North America. Concerning Northern Eurasia and the Middle East, the two associated gridded data sets have a limited overlapping with the Euro-Mediterranean data set. Further details as well as maps of the covered regions can be found in the associated publications [*Higgins et al.*, 2000; *Jones et al.*, 2009; *Haylock et al.*, 2008; *Yatagai et al.*, 2012] and the related web-sites (Table S2).

Since the eight GCMs have different resolutions and grids, and precipitation is highly dependent on the spatial scale, all daily gridded values were remapped onto a common grid with the coarsest resolution of 1.5° by applying a conservative remapping procedure [*Chen and Knutson*, 2008]. By applying the same procedure, in order to allow for a comparison between gridded observations and model simulations for the historical period 1966-2005, observations were remapped to the common grid of the models.

As for the characterisation of the extremes and their changes in the 21th century, we compare the two 40-year time periods 2020-2059 and 2060-2099 with the historical period. The length of the two periods ensures an adequate data amount for the statistical inference that is known to be data-demanding for extremes. In order to assess the goodness-of-fit of the statistical model (hereafter, reliability), stationary processes within each 40-year period were assumed.

The analysis was performed in the frame of Extreme Value Theory by applying a Peaks Over Threshold approach [*Davison and Smith*, 1990]. In this context, the distribution of excesses over a high threshold (here, set as the 90^{th} percentile) can be modelled by using the Generalised Pareto (GP) family, i.e.

$$H_{\sigma,\xi}(y) = 1 - \{1 + (\xi y/\sigma)\}^{-1/\xi}, \quad \xi \neq 0$$

$$H_{\sigma,\xi}(y) = 1 - \exp(-y/\sigma), \quad \xi = 0 \tag{1}$$

where $\sigma > 0$, $y \geq 0$ when $\xi \geq 0$ and $y \in [0, -\sigma/\xi]$ when $\xi < 0$. The two parameters σ and ξ are called scale and shape parameter, respectively. As soon as an estimation for both parameters is available ($\hat{\xi}$ and $\hat{\sigma}$), the return level z_R (i.e., the value that is expected to be exceeded on average once every R years, here 50) can be estimated by:

$$z_R = u + \hat{\sigma}\hat{\xi}^{-1}\left[(R\zeta_u)^{\hat{\xi}} - 1\right] \tag{2}$$

where u is the chosen threshold and ζ_u is the intensity of the Poisson process which is assumed to describe the occurrence of the excesses. Concerning the parameters estimation and in order to avoid numerical problems connected with optimisation procedures, we applied the Generalised Probability Weighted Moments method (GPWM, see the Appendix) [*Diebolt et al.*, 2007].

The goodness-of-fit of the estimation was tested by a modified Anderson-Darling statistic [*Luceño*, 2006, and references therein], i.e.,

$$A = n \int_{-\infty}^{\infty} [H(y) - F_n(y)]^2 \cdot [1 - H(y)]^{-1} \, dy \tag{3}$$

where n denotes the number of excesses, H is the assumed theoretical distribution (here, Generalised Pareto) and F_n is the empirical distribution function. Since the parameters

of the distribution H were not known, the asymptotic distribution of A and, thus, the critical values (at the 0.95 level) for the test were also unknown. As a consequence of applying the GPWM method, the covariance of the Gaussian process to which the integrand of A asymptotically converges cannot be approximated. Thus, the critical values for the test were obtained by using a bootstrap procedure [*Babu and Rao*, 2004].

In brief for the bootstrap procedure, let $y_1, \ldots, y_n$ be the excesses and $\hat{\sigma}$ and $\hat{\xi}$ the estimated parameters. Then, m additional samples (in this exercise, 1000) can be generated from $H_{\hat{\sigma},\hat{\xi}}$ and the A statistic can be computed m times by estimating the shape and the scale from the generated m samples. The critical values for the test can be derived from the calculated m values of A.

In order to perform an inter-model comparison with respect to observations in the historical period 1966-2005, Taylor diagrams were used for the estimated return level fields [*Taylor*, 2001].

3. Results

Changes in precipitation extremes are presented in terms of very high risk events, i.e., 50-year return levels (values that are expected to be exceeded on average once every 50 years) derived by the inferred distributions.

In the historical period, as shown in Figure 1, a reliable characterisation of daily extreme precipitation cannot be achieved for larger areas of the world, where an estimation of the return levels cannot be obtained. This is the case during boreal winter for a belt elongated over the subtropics and tropics of the NH and the oceanic areas west of the three continents of the SH. In boreal summer, unreliable estimations (i.e., failing the

goodness-of-fit test) are found mainly over the eastern North Pacific, north-eastern Africa and Arabian Peninsula as well as a large part of the Mediterranean basin and the eastern North Atlantic, eastern South Pacific and north-central Australia. A similar spatial pattern is also identified for spring and autumn (not shown). Lack of reliability roughly corresponds to areas characterised by higher positive values of the shape parameter (not shown) and therefore strongly heavy-tailed Generalised Pareto distributions. Although spatial differences do exist, this correspondence is a common feature of seven out of eight models and could be connected with the parameterisation of convection in regions receiving smaller amounts of seasonal precipitation [*Dai*, 2006], poorly represented land- and ocean-atmosphere interactions as well as deficiencies in the representation (position and shape) of the Intertropical Convergence Zone [*Huang et al.*, 2004; *Dai*, 2006; *Richter and Xie*, 2008; *Good et al.*, 2009].

For the mid and high latitudes, six out of eight models show a spatially homogeneous tail behaviour with slightly negative and positive values of the shape parameter (not shown). This means that the probability of precipitation extremes either has a finite upper bound or decreases approximately exponentially or slightly slower towards zero. Nevertheless, a glance at the individual simulations reveals remarkable inter-model differences as well as areas with a larger probability of higher extremes. In the Euro-Mediterranean area, northern Eurasia and North America, the simulations show lower inter-model variability and higher correlation with the observations in boreal winter (Figs. 1 and S1). Conversely, for Australia, southern Asia and the Middle-East all seasons are characterised by larger inter-model variability and lower correlation with the observations (Figs. S1-S2).

For the period 2020-2059, both scenarios reveal reliable and consistent changes only for scattered areas in the mid and high latitudes of both hemispheres (Figs. 2 and S3). A similar global pattern with regional differences is estimated for the other seasons (not shown). It is worth noting that the intensity reduction over the northern tropical Atlantic is strongly seasonally dependent as it almost disappears in boreal summer and is less pronounced in spring and autumn.

Towards the end of the 21st century (2060-2099), a similar pattern but with more pronounced changes compared to the middle of the century is projected under the RCP8.5 scenario. For the RCP4.5 scenario, for which the radiative forcing stabilises in the second half of the 21st century, changes in extremes are less pronounced. Consistent and reliable increases of precipitation extremes are obtained for all seasons over the mid and high latitudes of both hemispheres mainly for the RCP8.5 scenario. In the SH, the spatial pattern of consistent and reliable areas does not show a marked seasonal dependence. In the NH within the zone showing consistency and reliability, different areas can be highlighted for each season (potentially connected with sea ice changes [e.g. *Budikova*, 2009; *Screen et al.*, 2013]), for instance Northern Eurasia in boreal winter; the North Pacific, the northwestern Atlantic/Arctic Ocean in boreal summer (Fig. 2). Meridional differences are clearer in the zonal means (Figs. 3 and S4). They show more pronounced increases over the high-latitudes of both hemispheres in all seasons, with the exception of the NH in the mid-century boreal summer, associated with larger inter-model variability. Over the SH, a sharp decrease in the estimated positive changes from the high to the mid latitudes is evident in all seasons and, with the exception of the austral winter, followed by a strong

increase towards the low latitudes. Over the NH, the poleward meridional increase of the estimated positive changes is almost continuous in boreal winter (Fig. 3) and autumn (not shown), while a stepwise poleward increase is projected for summer (Fig. 3) and spring (not shown). Stronger hemispheric differences (Fig. 3) are estimated over the high latitudes for RCP8.5 at the second half of the century that are most prominent in summer (11% difference between the NH and SH spatial means) and autumn (15% difference between the NH and SH spatial means). The largestLargest changes at the second half of the century are found over the high latitudes of the NH for autumn (45%) and for spring in the SH (39%). For the mid latitudes in the same period, mean regional changes are highest in the NH autumn (37%) and the SH summer (30%). No reliable assessment can be made for the subtropical-tropical regions (Fig. 3). The identified increase of extremes for the 21st century (although seasonally and regionally dependent) is higher than previously estimated for annual extremes [*Kharin et al.*, 2013]. The effect of stabilisation of the radiative forcing in the RCP4.5 scenario is evident in Figure S4, showing less pronounced differences between the two periods 2020-2059 and 2060-2099 compared to the RCP8.5 scenario.

In order to gain a better insight into the regional changes, twenty six land areas [*IPCC*, 2012] were selected and the inter-model variability of the regional means is provided in the Figure 4 and Figs. S5-7. Remarkable seasonal and regional differences are evident among the twenty six land-areas. Reliable extremes characterisation can be made for 62% (65%) of the land-areas in boreal winter (summer) and for 88% (85%) of the areas in spring (autumn). For some areas, reliability shows a clear seasonal dependency (e.g.,

189 southern Europe), while this is not the case for regions such as Northern Europe and
190 Northern Asia. Finally, it is evident from the results that the regional averages show a
191 better agreement between models.

192

4. Discussion

193 In the tropics, the identified lack of reliability and consistency in extreme precipitation
194 could be associated with a deficiency in the representation of upward velocities that seems
195 to introduce large differences in climate models output, an underestimation of the response
196 to global warming [*Allan and Soden*, 2008; *O'Gorman and Schneider*, 2009] as well as
197 with model difficulties in reproducing processes based on organised convective systems
198 [*Zhang*, 2005; *Benedict and Randall*, 2007]. Conversely in the mid and high latitudes,
199 where large-scale processes play an important role [*O'Gorman and Schneider*, 2009], re-
200 liable and consistent results are coherent with the increase of precipitation extremes not
201 only constrained byare coherent with the Clausius-Clapeyron constraint equation but also
202 dynamically driven and potentially connected withand could be dynamically linked to the
203 polar warming amplification [*Schlosser et al.*, 2010; *Jaiser et al.*, 2012; *Francis and Vavrus*,
204 2012; *Screen and Simmonds*, 2013], although this connection needs to be investigated and
205 is our next research issue.
206 At the regional level, models show a better agreement on the projected increase of return
207 levels over land, although large variability affects the estimated seasonal changes over
208 specific areas (e.g., Eastern Asia in summer). Finally, it is worth to point that for some
209 areas such as the Indian Monsoon region, where models deficiencies were also identified

by *Hasson et al.* [2013] and *Sperber et al.* [2013], reliable estimations cannot be achieved.

Appendix A: the GPWM method

The GPWM for the Generalised Pareto distribution are defined by

$$\mu_\omega = E\left\{Y\omega\left(1 - H_{\sigma,\xi}(Y)\right)\right\} \tag{A1}$$

where E denotes the expected value and ω is a continuous function null (and with right derivative) at 0. An estimator of μ_ω is given by [*Diebolt et al.*, 2007]:

$$\hat{\mu}_{\omega,n} = \int_0^\infty W\left(1 - F_n(x)\right)\, dx \tag{A2}$$

where n represents the number of excesses, F_n denotes the empirical distribution function of the excesses and W is the primitive of ω. For applications (as in the current exercise) a good choice for the function ω is $\omega(x) = x^r$ with $r = 1, 1.5$. This implies that $\mu_\omega = \mu_r$ can be estimated by:

$$\hat{\mu}_r = n^{-1}\sum_{i=1}^{n} Y_{(i)} \cdot \left[(n-i)n^{-1}\right]^r \tag{A3}$$

where $Y_{(i)}$ represents the ordered sample of the excesses. Finally, the estimated parameters are provided by using the following equalities, replacing μ_1 and $\mu_{1.5}$ with their estimates:

$$\sigma = (2.5\mu_{1.5}\mu_1) \cdot (2\mu_1 - 2.5\mu_{1.5})^{-1} \tag{A4}$$

$$\xi = \left[4\mu_1 - (2.5)^2\mu_{1.5}\right] \cdot (2\mu_1 - 2.5\mu_{1.5})^{-1} \tag{A5}$$

This approach is valid for $\xi \in (-1, 1.5)$.

Acknowledgments. A.T., J.L., E.X. and P.N. acknowledge support from the EU-FP7 ACQWA project (n. 212250). A.T. acknowledges DFG (grant TO829/1-1). We thank F. Albrecht for computational support and P. Fraschetti for data retrieving and computational support. We acknowledge the World Climate Research Programmes Working Group on Coupled Modelling for model simulations. We thank EU-FP6 project ENSEMBLES, NOAA/OAR/ERL PSD, Aphrodite's project and the Bureau of Meteorology Australia for gridded observations. We thank two anonymous reviewers for their comments and suggestions.

References

Alexander, L. V., *et al.*, Global observed changes in daily climate extremes of temperature and precipitation *J. Geophys. Res.*, *111*, D05109.

Allan, R. P., and B. J. Soden (2008), Atmospheric Warming and the Amplification of Precipitation Extremes, *Science*, *321*, 1481–1484.

Babu, G. J., and C. R Rao (2004), Goodness-of-fit tests when parameters are estimated, *Sankhya*, *66*, 63–74.

Benedict, J. J., and D. A. (2007), Randall Observed characteristics of the MJO relative to maximum rainfall, *J. Atmos. Sci.*, *64*, 2332–2354.

Budikova, D. (2009), Role of Arctic sea ice in global atmospheric circulation: A review, *Global Planet. Change*, *68*, 149–163.

Chen, C. T., and T. Knutson (2008), On the Verification and Comparison of Extreme Rainfall Indices from Climate Models, *J. Climate*, *21*, 1605–1621.

Dai, A., (2006), Precipitation characteristics in eighteen coupled climate models, *J. Climate*, *19*, 4605–4630.

Davison, A. C., and R. L. Smith (1990), Models for exceedances over high thresholds, *J. Roy. Stat. Soc. B Met.*, *52*, 393–442.

Diebolt, J., A. Guillou, and I. Rached (2007), Approximation of the distribution of excesses through a generalized probability-weighted moments method, *J. Stat. Plan. Infer.*, *137*, 841–857.

Francis, J. A., and S. J. Vavrus (2012), Evidence linking Arctic amplification to extreme weather in mid-latitudes, *Geophys. Res. Lett.*, *39*, L06801.

Good, P., J. A. Lowe, and D. P. Rowell (2009), Understanding uncertainty in future projections for the tropical Atlantic: relationships with the unforced climate, *Clim. Dyn.*, *32*, 205–218.

Haylock, M. R., *et al.* (2008), A European daily high-resolution gridded dataset of surface temperature and precipitation, *J. Geophys. Res.*, *113*, D20119.

Hasson, S., V. Lucarini, and S. Pascale (2013), Hydrological cycle over south and southeast Asian river basins as simulated by PCMDI/CMIP3 experiment, *Earth. Sys. Dynam. Discuss.*, *4*, 109–177.

Higgins, R. W., *et al.* (2000), *Improved United States precipitation quality control system and analysis*, NCEP/Climate Prediction Center Atlas.

Hofstra, N., *et al.* (2009), Testing E-OBS European high-resolution gridded dataset of daily precipitation and surface temperature, *J. Geophys. Res.*, *114*, D21101.

Huang, B., P. S. Schopf, and J. Shukla (2004), Intrinsic ocean-atmosphere variability of the tropical atlantic ocean, *J. Climate*, *17*, 2058–2077.

IPCC (2012), *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*, C. B. Field et al. Eds, Cambridge University Press, New York.

Jaiser, R., *et al.* (2012), Impact of sea ice cover changes on the Northern Hemisphere atmospheric winter circulation, *Tellus A*, *64*, 11595.

Jones, D. A., W. Wang, and R. Fawcett (2009), High-quality spatial climate data-sets for Australia, *Aust. Meteorol. Ocean. J.*, *58*, 233–248.

Kharin, V. V., *et al.* (2013), Changes in temperature and precipitation extremes in the CMIP5 ensemble, *Climatic Change*, doi 10.1007/s10584-013-0705-8.

Luceño, A., (2006), Fitting the generalized Pareto distribution to data using maximum goodness-of-fit estimators, *Comput. Stat. Data An.*, *51*, 904–917.

Min, S. K., *et al.* (2011), Human contribution to more-intense precipitation extremes, *Nature*, *470*, 378–381.

Moss, R. H., *et al.* (2010), The next generation of scenarios for climate change research and assessment, *Nature*, *463*, 747–756.

O'Gorman, P. A., and T. Schneider (2009), The physical basis for increases in precipitation extremes in simulations of 21st-century climate change, *Proc. Nat. Acad. Sci. U.S.A.*, *106*, 14773–14777.

Richter, I., and S. P. Xie (2008), On the origin of equatorial Atlantic biases in coupled general circulation models, *Clim. Dyn.*, *31*, 587–598.

Screen, J. A., and I. Simmonds (2013), Exploring links between Arctic amplification and mid-latitude weather, *Geophys. Res. Lett.*, *40*,959–964.

Screen, J. A., *et al.* (2013), Atmospheric impacts of Arctic sea-ice loss, 19792009: separating forced change from atmospheric internal variability, *Clim. Dyn.*, doi:10.1007/s00382-013-1830-9.

Schlosser, E., *et al.* (2010), Characteristics of high precipitation events in Dronning Maud Land, Antarctica, *J. Geophys. Res.*, *115*, D14107.

Scoccimarro, E., *et al.* (2013), Heavy precipitation events in a warmer climate: results from CMIP5 models, *J. Climate*, doi:10.1175/JCLI-D-12-00850.1.

Sperber, K. R, *et al.* (2013), The Asian summer monsoon: an intercomparison of CMIP5 vs. CMIP3 simulations of the late 20th century, *Clim. Dyn.*, doi:10.1007/s00382-012-1607-6.

Sun, Y., *et al.* (2006), How Often Does It Rain?, *J. Climate*, *19*, 916–934.

Taylor, K. E., (2001), Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, *106*, 7183–7192.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, *B. Am. Meteorol. Soc.*, *93*, 485–498.

Trenberth, K. E., *et al.* (2007), in *Climate change 2007. The physical science basis. Intergovernmental Panel on Climate Change 4th assessment report*, S. Solomon et al. Eds, Cambridge University Press, New York, 235–336.

312 Trenberth, K. E. (2011), Changes in precipitation with climate change, *Clim. Res.*, *47*,
313 123–138.
314 Westra, S., L. V. Alexander, and F. W. Zwiers (2013), Global Increasing Trends in Annual
315 Maximum Daily Precipitation, *J. Climate*, *26*, 3904–3918.
316 Yatagai, A., *et al.* (2012), APHRODITE Constructing a long-term daily gridded precipi-
317 tation dataset for Asia based on a dense network of rain gauges, *B. Am. Meteorol. Soc.*,
318 *93*, 1401–1415.
319 Zhang, C., (2005), Madden-Julian Oscillation, *Rev. Geophys.*, *43*, RG2003.
320 Zhang, X., *et al.* (2007), Detection of human influence on twentieth-century precipitation
321 trends, *Nature*, *448*, 461–466.

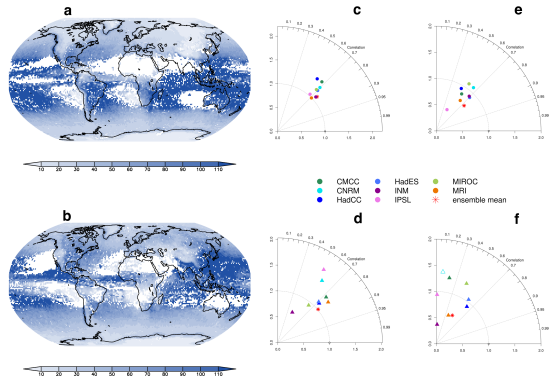

Figure 1. Ensemble mean 50-year return levels (mm) estimated for the period 1966-2005 in boreal winter **a** and summer **b**. Blue coloured areas identify grid points where at least 75% of the models pass the goodness-of-fit test (reliable points). Taylor diagrams for estimated 50-year return levels in winter and summer over Northern Eurasia **c,d** and North America **e,f**. The full symbols denote models with at least 75% of reliable grid points in the region.

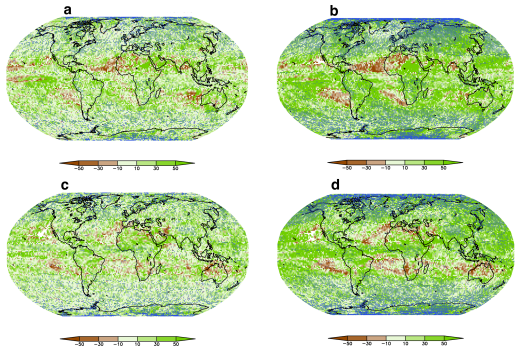

Figure 2. Ensemble mean changes of the estimated 50-year return levels (%) with respect to the period 1966-2005, under the RCP8.5 scenario for winter 2020-2059 **a** and 2060-2099 **b** and summer 2020-2059 **c** and 2060-2099 **d**. Blue dots mark grid points where at least 75% of the models pass the goodness-of-fit test and agree on the sign of the estimated changes.

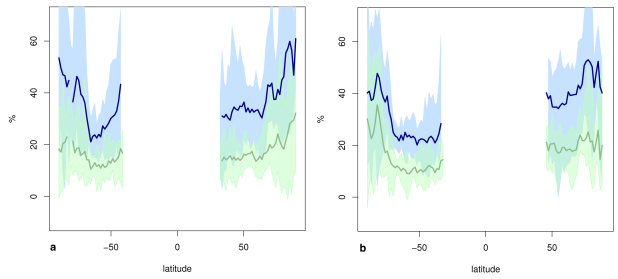

Figure 3. Zonal mean changes of the estimated 50-year return levels (%) with respect to the period 1966-2005 in winter **a** and summer **b** under the RCP8.5 scenario. Blue and green lines represent the ensemble mean for the periods 2020-2059 and 2060-2099, respectively. Blue and green shaded areas show the inter-model variability for the periods 2020-2059 and 2060-2099, respectively. The ensemble mean and the inter-model variability are plotted only when at least six models out of eight provide at least 75% of reliable grid points for the zonal mean.

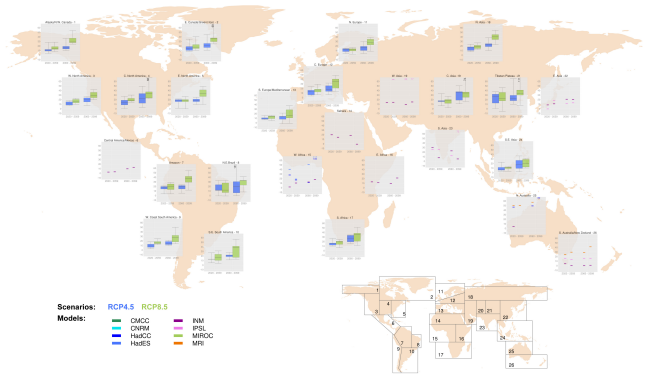

Figure 4. Regional mean changes of winter 50-year return levels (%) in the periods 2020-2059 and 2060-2099 (RCP4.5 and RCP8.5 scenarios) with respect to the reference 1966-2005. Regional means are derived for each model using only land grid points. Box-plots are drawn when at least four models contribute with a minimum of 75% of reliable land points in the region. The black line is the median of the regional means; the whiskers of the box-plots represent the maximum and the minimum of the areal means. Point-plots are drawn when the above conditions are not satisfied.