Weather extremes over Europe under 1.5°C and 2.0°C global warming from HAPPI regional climate ensemble simulations

Kevin Sieck¹, Christine Nam¹, Laurens M. Bouwer¹, Diana Rechid¹, Daniela Jacob¹
¹Climate Service Center Germany (GERICS), Helmholtz-Zentrum Geesthacht, Hamburg, 20095, Germany

Abstract. This paper presents a novel data set of regional climate model simulations over Europe that significantly improves our ability to detect changes in weather extremes under low and moderate levels of global warming. The data set provides a unique and physically consistent data set, as it is derived from a large ensemble of regional climate model simulations. These simulations were driven by two global climate models from the international HAPPI consortium. The set consists of 100 x 10-year simulations and 25 x 10-year simulations, respectively. These large ensembles allow for regional climate change and weather extremes to be investigated with an improved signal-to-noise ratio compared to previous climate simulations. The changes in four climate indices for temperature targets of 1.5°C and 2.0°C global warming are quantified: number of days per year with daily mean near-surface apparent temperature of >28°C (ATG28); the yearly maximum 5-day sum of precipitation (RX5day); the daily precipitation intensity of the 50-yr return period (RI50yr); and the annual Consecutive Dry Days (CDD).

This work shows that even for a small signal in projected global mean temperature, changes of extreme temperature and precipitation indices can be robustly estimated. For temperature related indices changes in percentiles can also be estimated with high confidence. Such data can form the basis for tailor-made climate information that can aid adaptive measures at a policy-relevant scales, indicating potential impacts at low levels of global warming at steps of 0.5°C.

1 Introduction

Identifying regional climate change impacts differ for different global mean temperature targets is increasingly relevant to both the private sector, as investors demand financial disclosure associated with climate change risks and opportunities (Goldstein, et al., 2018), as well as the public sector, as national climate action policies are developed. This is especially true after the adoption of the Paris Agreement of the United Nations, which aims to keep global climate warming well below 2.0°C compared to pre-industrial times (UNFCCC, 2015). Temperature targets, however, are not directly related to the representative concentration pathways (Van Vuuren et al., 2011) used in the current generation global climate simulations (CMIP5, Taylor et al., 2012). Therefore, new techniques are being developed to extract information on the possible implications of further global warming. Recent studies using CMIP5 data have shown that climate change indices can be extracted for different warming levels, by identifying specific time periods when a certain global mean temperature (GMT) warming is reached in a general circulation model (GCM) (Schleussner et al., 2016; Vautard et al., 2014; Jacob et al., 2018). These studies typically
use the 5 to 15 ensemble members available in CMIP5 for their global and regional studies. Mitchell et al. (2016) argue however that a different experiment design is needed to better address the policy-relevant temperature targets with climate simulations, because the relatively small CMIP5 ensemble does not provide the necessary size to quantify changes in weather extremes at low levels of warming. Indeed, the high natural variability in models requires the creation of large ensemble datasets (Deser et al., 2013). Following the recommendations of Mitchell et al. (2016), the HAPPI consortium (“Half a degree Additional warming, Prognosis and Projected Impacts”) designed targeted experiments created for the purpose of extracting the required information on distinct warming levels using 10 state-of-the-art GCMs (Mitchell et al., 2017). The HAPPI experiments include a large number of ensemble members, typically 50 to 100 members per GCM, using AMIP-style integrations (Gates et al., 1992), which significantly improves the signal-to-noise ratios. A better signal-to-noise ratio is essential for differentiating between impacts from 1.5°C and 2.0°C global warming, especially for changes in weather extremes.

To bridge the gap between GCM model output and regional climate impact assessments, which require a much higher resolution than GCMs (Giorgi and Jones, 2009), the Regional Climate Model (RCM) REMO (Jacob et al., 2012) is used to dynamically downscale simulations from two GCMs from the HAPPI consortium. Dynamical downscaling with RCMs is one option to bridge the gap between GCM model output and regional climate impact assessments which provides physically consistent high-resolution climate information (Jacob et al., 2014; Giorgi and Gutowski, 2015; Gutowski et al., 2016). Some other recent studies are also successfully pursued the creation of large ensemble datasets on the basis regional climate models, of up to 50 members, to study amongst others rainfall extremes (Leduc et al., 2019). Here, we develop two regional climate datasets of 25 and 100 members. To demonstrate the potential of this data set for regional climate impact studies, under 1.5°C and 2.0°C global warming, changes in four climate indices for weather extremes are quantified.

In Section 2, we present the REMO regional climate model, experiment setup and simulations performed. In Section 3, four relevant climate indices for extreme weather which can be derived from the HAPPI data set are presented, and lastly, conclusions are derived in Section 4.

2 Methods

2.1 Model simulations

To create a data set for regional climate impact studies for Europe under 1.5°C and 2.0°C global warming the regional climate model REMO was dynamically downscaled using the HAPPI GCM simulations. All the HAPPI GCM simulations use sea-surface temperatures (SST) for respective periods, known as AMIP-style (Gates et al., 1992), representing the following three periods: A current decade (2006-2015) with observed SSTs, and two projected periods with 1.5°C and 2.0°C warmer (global mean surface temperature) than pre-industrial (1861-1880) conditions. For the two warmer periods, CMIP5 mean SST anomaly patterns for the respective global warming are added to the observed SST pattern used for the current decade. Lastly,
greenhouse gas forcing is constructed from RCP2.6 and RCP4.5 emission scenarios, respectively. A comprehensive description of the HAPPI experiment design is given in Mitchell et al. (2017).

The RCM REMO is a hydrostatic limited-area model of the atmosphere that has been extensively used and tested in climate change studies over Europe (Jacob et al., 2012; Teichmann et al., 2013; Kotlarski et al., 2014). The simulation domain follows the CORDEX specification for the standard European domain with 0.44° horizontal resolution. To exclude the sponge zone, where the REMO simulations are relaxed towards the GCM solutions, from the core domain defined by CORDEX the entire domain has 121x129 grid boxes. In the vertical 27 levels are used without any nudging except for the boundaries. Boundary conditions are taken from the HAPPI Tier1 experiments (Mitchell et al., 2017), which are carried out with ECHAM6 (Stevens et al., 2013) (100 members per period) and NorESM (Bentsen et al., 2013) (25 members per period) that provide 6-hourly 3-dimensional data for downscaling. In REMO the same green-house-gas forcings as for the GCMs were used. The SST was taken directly from the GCM output matching the GCM land-sea mask, which provides more consistent boundary conditions for the downscaling.

For each GCM member only one REMO simulation was carried out, as inter-member variability of an RCM ensemble over Europe on these time scales is small compared to the internal variability of a GCM (Sieck et al., 2016). Each simulation covers a period of ten years, and as such, initial conditions for the lower boundary need to be in balance with the RCMs internal climate in order to avoid artificial drifts in the modelled results. To achieve this, for each driving GCM, the first year of a random GCM member was simulated five times with REMO using initial conditions from the end of the previous run, creating one initial soil temperature state for every ensemble member in one period. This was performed for each of the three time periods. Tests showed that this minimizes drifts in the deep soil climatology compared to initial conditions taken directly from the GCM (not shown).

We are aware of the incorrect Sea-Ice concentrations used in ECHAM6. Due to the interpolation procedure for the Sea-Ice extent, it could happen that Sea-Ice was artificially created where no ice conditions were present in the original data-set, e.g., in summer in the Baltic Sea. ECHAM6 has a mechanism that as soon as there is a fraction of sea-ice greater than zero, the SST is limited to a maximum of 272.5K. This leads to artificial temperature jumps in the SST between adjacent grid boxes as soon as erroneous sea-ice appeared in one of the grid boxes. This error was inherited by the first set of REMO simulations and was corrected by using the originally provided SST fields from the HAPPI project for the current set. After testing different temperature and/or sea-ice fraction thresholds the authors decided to keep the original Sea-Ice maps, because in the cases where artificial sea-ice was created the fraction was usually well below 1% only in rare cases reaching up to 4% (not shown). All other settings would have removed too much sea-ice in other seasons or led to unrealistic gradients of sea-ice fraction.

With the tile approach of REMO the effect of the artificial sea-ice is hardly detectable even in the averaged near-surface variables.
2.2 Climate indices

To demonstrate how adaptation-relevant information can be derived from the HAPPI data set for two different temperature targets, four climate indices used in climate impact studies are presented. The extremes are selected based on recommend indices developed by the joint CCI/CLIVAR/JCOMM Expert Team on Climate Change Detection and Indices (ETCCDI) (Karl et al., 1999; Frich et al., 2002) and other indicators. The selected climate indices are:

- Number of days per year with a daily mean near-surface apparent temperature of more than 28°C (ATG28);
- Annual maximum 5-day sum of precipitation (RX5day);
- Relative change in daily precipitation intensity at the 50-yr return period (RI50yr);
- Consecutive Dry Days (CDD) as a measure of meteorological drought.

All four climate indices are calculated from the daily mean precipitation, temperature, and/or dew-point temperature output of the model; for each year and ensemble member. The ECHAM6 driven ensembles yield 1000 data points for each grid box and simulation period, and NorESM has 250 data points, respectively.

The ATG28 index is used as an indicator for heat stress, which is relevant for impacts on human health (Davis et al., 2016). The apparent temperature is computed using the same formulation as in Davis et al. (2016):

\[
AT = -2.653 + 0.994T + 0.0153T_d^2,
\]

with \(AT\) being the apparent temperature, \(T\) the daily mean near-surface temperature, and \(T_d\) the daily mean near-surface dewpoint temperature. Similar formulations exist in the literature showing very similar results (see Anderson et al., 2013 for a review). The threshold of 28°C is based on the definition of Zhao et al. (2015) who set this limit as the lower boundary for human heat stress.

The index for the annual sum maximum of the five-day precipitation sum (RX5day) is used to characterise heavy precipitation events, which can be relevant for flood generation in river basins. The RX5day represents a noisy, i.e., highly spatially and temporally variable parameter. A large ensemble would allow for a better assessment of the signal-to-noise in extreme precipitation.

A change in extreme precipitation directly influences local communities, through design standards chosen in order for structures to withstand a flood with a specified return period. As such, a Gumbel Type I extreme value distribution is fitted to the annual maxima of daily rainfall amounts. Using this distribution, an estimate is made of the intensity of rainfall events associated with a given exceedance probability. For each ensemble, the daily rainfall intensity for the 50-year return is computed, hereafter called RI50yr. Such information is useful for infrastructure design and maintenance. For example, road authorities in Europe typically use between 1 and 10-year return periods for assessing effects of rainwater falling on major roads (highways), and between 100 and 100 years for rainfall beside the road and waters crossing the road (Bless et al., 2018).
Lastly, the Consecutive Dry Days (CDD), is calculated for each of the PRUDENCE regions (Christensen et al., 2007). These regions are illustrated in Figure 1. The CDD is the maximum number of consecutive days with daily precipitation amount of less than 1 mm over a region (Karl et al., 1999; Peterson et al., 2001).

2.3 Changes in climate indices

The differences between the historical and temperature target 1.5°C and 2.0°C simulations for each of the climate indices are computed as follows: In case of ATG28, differences of the 5th, 50th, and 95th percentiles were computed by subtracting the ensemble mean of the current decade from the projected periods. Only areas with more than 20 non-zero data points in the reference period were included in the analysis of ATG28 in order to allow for confidence interval calculations for the shown percentiles using order statistics. Statistical significance is determined when the calculated percentile in the warming period is outside the percentile confidence range of the current period. As ATG28 is temperature-based, changes over the ocean surfaces are masked out, because they are to a large extent determined by the prescribed SST changes.
Differences for RX5day are computed by subtracting the ensemble mean of the current decade from the projected periods, similar to the ATG28, however the statistical significance for RX5day was calculated using a Mann-Whitney-U-test and only results are shown with a significance at the 95% level.

The differences in RI50yr are computed as the relative change in daily precipitation intensity of the 50-yr return period between the 1.5°C and 2.0°C simulations compared to the historical simulations, for the NOResm and ECHAM6-driven REMO simulations, respectively. Similar to the ATG28 analysis, the differences in the distribution of CDD are calculated with a Mann-Whitney U-Test with a significance at the 95% level, determining whether samples from the two periods are drawn from a population with the same distribution.

3. Results

3.1 Apparent temperature

Figures 2 and 3 show the changes in ATG28 for the NorESM and ECHAM6 driven ensembles, respectively. In general, the changes are strongest close to warm ocean areas, especially around the Mediterranean. But also the central and eastern parts of Europe show increases in ATG28, consistent with the increase in mean temperature. The distinct difference between the two warming levels should be noticed. In the 1.5°C period the increase in ATG28 is mostly moderate with up to 9 days whereas changes in the 2.0°C period are reaching 18 days and more in the median around the Mediterranean. This result is consistent between the ECHAM6 and NorESM driven ensembles. Also the changes across the percentiles are consistent, i.e., there is no change in the shape of the distribution of ATG28. It should be noted that the spatial resolution of the simulations allows to show the lower level of warming in mountainous areas compared to coastal areas, e.g., over Italy. This is especially important in areas with complex topography such as the Mediterranean, which is usually only poorly resolved in GCM simulations.
Figure 2. Differences in ATG28 between the current and the 1.5°C period (top row) respectively the 2.0°C period (bottom row) for the NorESM driven REMO simulations in number-of-days. Shown are the Differences in the 5th percentile (left column), median (middle column) and 95th percentile (right column). Differences over Ocean areas are masked out in grey, because they are closely related to the prescribed SST changes.
3.2 Five-day precipitation sum

Figure 4 shows the relative differences of RX5day for the four REMO ensemble experiments. In general, there is an increase in RX5day over the European part of the domain with stronger signals in the 2.0°C compared to the 1.5°C period. It can also be seen that the patterns in the ECHAM6 driven simulations are more coherent with larger areas showing a significant change. This is related to the difference in ensemble size and underlines the necessity for a large ensemble to achieve proper signal-to-noise ratios when looking at the difference in regional changes under small GMT increases in highly variable quantities such as precipitation extremes. Tests with a randomly picked 25-member ensemble from the ECHAM6 driven simulations showed a similar noisy pattern as the NorESM driven runs (not shown).

Apart from effects at the boundaries, the strongest signal in the interior of the simulation domain appears over the Baltic Sea, with an increase of up to 15% in RX5day under a 2.0°C increase in GMT. This result is consistent between both ensembles. A similar increase can be seen over the Adriatic Sea, but is not so pronounced in the ECHAM6 driven ensemble. This might be related to feedbacks from the unrealistic SSTs, because the GCMs usually do not resolve these small basins. In these locations the SST is interpolated from the nearest SST value of the GCM, which might not be adequate for the region.
Figure 4. Relative difference of RX5day (in percent) between current and the 1.5°C period (left column) respectively the 2.0°C period (right column) for the NorESM with 25 members (top row) and ECHAM6 with 100 members (bottom row) driven REMO simulations.

3.3 Daily rainfall intensity, 50-year return period

To account for the spatial differences in 50-year return period across Europe, the relative change (in percent) in daily rainfall intensity are presented in Figure 5. In the both the NorESM and ECHAM6 driven ensembles, a greater increase in the rainfall intensity is found in the 2.0°C simulations compared to 1.5°C. ECHAM6 driven simulations clearly show increases in the 24-hour rainfall intensity of the 50-yr return period, of up to 20% over continental Europe. The estimated changes in rainfall intensity in the NorESM driven simulations appear to be more extreme but these simulations are also more noisy as they are based on fewer ensemble members.
Figure 5. Relative difference of RI50yr (in percent) between current and the 1.5°C period (left column) respectively the 2.0°C period (right column) for the NorESM with 25 members (top row) and ECHAM6 with 100 members (bottom row) driven REMO simulations in %. 

Figure 5. Relative difference of RI50yr (in percent) between current and the 1.5°C period (left column) respectively the 2.0°C period (right column) for the NorESM with 25 members (top row) and ECHAM6 with 100 members (bottom row) driven REMO simulations in %.
3.4 Consecutive dry days

In this section, the changes in the Consecutive Dry Days (CDD) distributions of each Prudence regions of the two ensembles are compared. For each ensemble, the p-values of each region is presented in Table 1. Both the distributions 1.5°C and 2.0°C are compared to the historical CDD distribution, respectively. Where the p-value is greater or equal to a significance level, alpha, of 0.05 or smaller, the null hypothesis is rejected indicating the distributions differ. Bold numbers in Table 1 indicate that the distributions differ according to the test, whereas the non-bold numbers indicate that there is no statistical difference between the distributions.

<table>
<thead>
<tr>
<th>PRUDENCE Region</th>
<th>ECHAM6</th>
<th>NorESM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hist vs 1.5°C</td>
<td>Hist vs 2.0°C</td>
</tr>
<tr>
<td>1. (BI)</td>
<td>0.270</td>
<td>0.035</td>
</tr>
<tr>
<td>2. (IP)</td>
<td>0.036</td>
<td>0.000</td>
</tr>
<tr>
<td>3. (FR)</td>
<td>0.391</td>
<td>0.230</td>
</tr>
<tr>
<td>4. (ME)</td>
<td>0.077</td>
<td>0.015</td>
</tr>
<tr>
<td>5. (SC)</td>
<td>0.238</td>
<td>0.356</td>
</tr>
<tr>
<td>6. (AL)</td>
<td>0.333</td>
<td>0.105</td>
</tr>
<tr>
<td>7. (MD)</td>
<td>0.069</td>
<td>0.036</td>
</tr>
<tr>
<td>8. (EA)</td>
<td>0.325</td>
<td>0.465</td>
</tr>
</tbody>
</table>

Table 1. Mann-Whitney U-Test p-values for distributions of Consecutive Dry Days (CDD) for the ECHAM6 and NorESM driven ensembles for different PRUDENCE regions (for locations see Figure 1).

We begin by looking at three regions where the Mann-Whitney U-Test provided consistent results across the ensembles. In region 2, the Iberian Peninsula, the CDD distributions in both the 1.5°C vs. 2.0°C simulations differ statistically compared to the historical simulations. The substantial shift towards longer periods of dry days over this region can be seen in Figure 6. In contrast, regions 6 and 8, the Alps and Eastern European region, have CDD distributions which are statistically indistinguishable in the simulations under additional global warming compared to the historical simulation (Table 1). One can deduce that region 2, will suffer from more frequent and longer drought periods than experienced before compared to regions 6 and 8. Interestingly, for region 7, the Mediterranean, the CDD distributions of the ECHAM6 and NorESM ensembles of 1.5°C do not differ statistically from the historical period, yet both ensembles show a statistically different distribution at 2.0°C. Thus, one can conclude for region 7, according to these simulations, a global target of 1.5°C vs. 2.0°C increase in GMT will have an impact, motivating adaption and mitigation practices. In region 3, France, the results of the two ensembles differ. The ECHAM6 simulations suggest there is no difference in CDD distributions between the warmer climate compared to the historical period; whereas the NorESM simulations find the warmer climate has a distinctly different CDD distribution.
compared to the historical period. For Regions 1, 4, and 5 the Mann-Whitney U-test yields differing results between the ensembles and therefore results are not displayed in Figure 6.

Figure 6. Duration of drought events in three PRUDENCE regions (2=Iberian Peninsula, 3=France, 7=Mediterranean) under 1.5° and 2.0°C global warming. For significance see Table 1.

4. Discussion and conclusions

A unique climate data set has been presented that enables the quantification of differences between a 1.5°C and 2.0°C warmer world compared to pre-industrial times on a regional level. This data set can support climate change impact studies on the regional scale with physically consistent data, which is often not possible to achieve with other methods than dynamical downscaling. The use of a large ensemble (10 x 100 years) compared to alternative data sets for analysing changes under different temperature targets is especially beneficial to assess changes in highly variable meteorological parameters, such as extreme temperature and precipitation.
Here the 100 members driven by ECHAM6 provides information of statistically significant changes over relatively large and spatially homogeneous areas. In comparison, the 25-member ensemble driven by NorESM shows a much noisier spatial pattern which lowers confidence in the projected changes.

The significant differences in apparent temperature ATG28 under different global mean warming level show that a 0.5°C larger global mean warming can have considerable consequences for human health. This is especially true around the Mediterranean, where changes towards more hot and humid conditions along the coasts can have negative impacts on the population and may increase mortality due to heat stress. The tourism sector may also be negatively affected by hotter and more humid conditions. Robust estimates of percentiles and changes in percentiles can be derived from the large ensemble.

The yearly maximum 5-day sum of precipitation RX5day shows a general increase over Europe which is more pronounced under higher global mean warming. More coherent spatial pattern with larger areas showing significant changes result from the larger ensemble driven by ECHAM6 (100 members) compared to the smaller ensemble driven by NorESM (25 members) and also compared to a smaller sub-ensemble driven by ECHAM6, which underlines the need for big ensemble size to reliably detect changes in highly variable quantities such precipitation extremes.

With regard to the relative change in the daily rainfall intensity at the 50-year return period, a greater increase in rainfall intensity was found in the 2.0°C warmer world. Given these changes, information can be derived for local communities, which must consider changes in rainfall intensity when designing hydraulic and water resource infrastructures, as well as transportation infrastructure, including highways and bridges. Cost considerations associated with increasing rain intensity demands can be computed for upcoming design projects to ensure investments remain beneficial. The HAPPI data set can be used to calculate other return periods, catering to the demands of individual sectors.

Robust high and low percentile changes for precipitation are still difficult to distil on a grid box level from the data because of the high variability of precipitation extremes, but methods such as spatial aggregation might help to achieve robust signals on larger spatial scale.

The changes to Consecutive Dry Day distributions show that Spain will have drought conditions unlike what they have experienced in the historical period, even at a 1.5°C increase in GMT. For Italy, droughts associated with the 1.5°C simulations are historically similar, yet droughts associated with 2.0°C would be statistically unlike what has been experienced in the pre-industrial period, thus demonstrating the consequences of exceeding the 1.5°C GMT target of the Paris agreement.

The current data set was created using the only two GCMs available at the time for downscaling, one with a reduced number of ensemble members. As more GCM ensembles become available for downscaling in the future, it will allow for new studies which can provide more robust estimates of inter-model variability/uncertainty. Nevertheless, there is currently a unique data set targeted to the Paris agreement goals available for further analysis. Future plans include the creation of a similar regional data set for Africa. A comparison with alternative methods for extracting the warming level is lacking and should be done in future studies.
Data availability. The datasets generated and analysed for this study can soon be found in the long-term archive of DKRZ:
http://cera-www.dkrz.de/

Author contributions. KS and DJ designed the study. KS performed the regional climate model simulations. KS and CN did the analysis of the climate indices and created the plots. KS, CN, LMB and DR wrote the initial draft paper. All Authors contributed with discussions and revisions.

Competing interests. The authors declare that they have no conflicts of interest.

Acknowledgements. The authors like to thank German Climate Computing Center (DKRZ) for supplying computing time and the HAPPI consortium for providing the boundary conditions for the downscaling. The authors thank their co-workers for the valuable discussions. This work is part of the research project HAPPI-DE and the authors acknowledge the support from the German Federal Ministry of Education and Research (BMBF) for this project under grant number 01LS1613C. It was also supported by the Digital Earth project of the Helmholtz Association.

References

Christensen, J. H., Christensen, O. B.: A summary of the PRUDENCE model projections of changes in European climate by the end of this century, Climatic Change, 81, 7-30, 2007.

14


Preprint. Discussion started: 20 February 2020 © Author(s) 2020. CC BY 4.0 License.


