Robustness of Ensemble Climate Projections Analyzed with Climate Signal Maps: Seasonal and Extreme Precipitation for Germany

Susanne Pfeifer 1,*, Katharina Bülow 1, Andreas Gobiet 2,†, Andreas Hänsler 1, Manfred Mudelsee 3, Juliane Otto 1, Diana Rechid 1, Claas Teichmann 1 and Daniela Jacob 1

Abstract: Climate signal maps can be used to identify regions where robust climate changes can be derived from an ensemble of climate change simulations. Here, robustness is defined as a combination of model agreement and the significance of the individual model projections. Climate signal maps do not show all information available from the model ensemble, but give a condensed view in order to be useful for non-climate scientists who have to assess climate change impact during the course of their work. Three different ensembles of regional climate projections have been analyzed regarding changes of seasonal mean and extreme precipitation (defined as the number of days exceeding the 95th percentile threshold of daily precipitation) for Germany, using climate signal maps. Although the models used and the scenario assumptions differ for the three ensembles (representative concentration pathway (RCP) 4.5 vs. RCP8.5 vs. A1B), some similarities in the projections of future seasonal and extreme precipitation can be seen. For the winter season, both mean and
extreme precipitation are projected to increase. The strength, robustness and regional pattern of this increase, however, depends on the ensemble. For summer, a robust decrease of mean precipitation can be detected only for small regions in southwestern Germany and only from two of the three ensembles, whereas none of them projects a robust increase of summer extreme precipitation.

**Keywords:** regional climate models; climate change; robustness; precipitation; Germany; EURO-CORDEX; ENSEMBLES

1. Introduction

For climate change impact assessment, as well as for the design and evaluation of climate change adaptation measures, spatially high-resolution information on potential future regional climate changes is needed (e.g., [1,2]). This information can be provided by regional climate models (RCMs), which are used to refine the climate change information provided by global climate models (GCMs) on a much coarser spatial resolution. The GCMs are fed by potential future evolutions of concentrations of greenhouse gases and aerosols (or emissions in the case of recent Earth system models). One combination of emission scenario, GCM and RCM, results in one single possible future climate evolution for a region. Using a different GCM, a different RCM or assuming a different emission scenario gives a second possible future, and so on. However, it is not possible to decide which one is more likely.

There are three main sources contributing to this uncertainty. The first one is the unknown development of future anthropogenic greenhouse gas emissions and the partly unknown future storage capacities of the biosphere and the oceans. A second contribution to the uncertainty is the models themselves, which do not represent all processes of the climate system perfectly and which might represent different processes unequally well. Thirdly, the non-linearity of the processes constituting the climate system creates non-deterministic variability and prevents a clear identification of changing climate conditions from the climate simulations. These sources of uncertainty of climate projections have been described and analyzed in detail (e.g., [3,4]).

Some aspects of the uncertainty of climate projections have the potential to be reduced. A very important approach to address the uncertainty of climate projections is the improvement of process understanding and, thereby, the reduction of model biases [3]. In addition, increasing the spatial and temporal resolution of the simulations can lead to better representation of today’s climate by the models, including a more realistic representation of the variability and extreme event statistics [5,6]. This model improvement is the basis for the applied usage of climate simulation data. However, it cannot reduce the uncertainties related to the unknown future greenhouse gas emissions and to the natural internal variability of the climate system. It was shown that large parts of the uncertainty range (defined by two standard deviations) of the CMIP5 [7] projections can be covered by the uncertainty range of an initial condition ensemble of one single model [8]. This is especially true for the upcoming decades [3,9]. Decisions on adaptation to climate change are thus to be taken based on information provided by multi-model, multi-scenario and multi-realization ensembles of climate change simulations. For some climatic parameters, this information is (relatively) stringent, where the ensemble of simulations is
showing a clear trend despite the possible bandwidth in the magnitude of the trend. For other climatic parameters, the ensemble results in less clear information, with disagreement not only in the magnitude, but also in the direction of the projected changes.

To assess the results of ensembles of climate change projections and to display not only mean projected climate changes, but also information on the robustness of the results, various approaches of ensemble compositions, model output statistics and visualization methods have been developed. In the context of the European FP7 project IMPACT2C, robust changes in mean and extreme temperature, precipitation, winds and surface energy budgets were identified for Europe under 2 °C global warming, using a simple metric of model agreement on the sign of the changes as indicator of robustness [10]. A complex measure of robustness using cumulative density functions of the individual and multi-model mean projections to distinguish the signal from noise was applied to the global CMIP5 simulations. The authors concluded that the local spread of the simulations did not change substantially between CMIP3 and CMIP5 simulations [11]. To overcome the large uncertainties in the projections of local extremes, a method of spatial aggregation was applied to an initial condition ensemble of one Earth system model. Despite large uncertainties on the local and short-term scale, this method shows that projections of extremes are remarkably consistent from an aggregated spatial probability perspective [8].

The summary for policymakers of the 5th Assessment Report of the Intergovernmental Panel on Climate Change presents mean projected climate changes from an ensemble of global climate change simulations. The results are presented in a figure, including stippling of regions where model agreement is large and where the natural internal variability is small compared to the the multi-model mean change and hatching of regions where the natural internal variability is large compared to the multi-model mean change (Figure SPM-08a,b [12,13]). These figures comprise plenty of information, and this complexity makes it difficult to gather all information at once.

The aim of the present study is to reduce the complexity of such figures, in order to make them understandable, not only for climate experts, but also for decision makers from public administrations and from the economy. Here, we present the method of climate signal maps to derive and display the robustness of climate changes projected by an ensemble of regional climate change simulations. The method has similarities to the IPCC method described above, but gives simplified information understandable at a glance. The paper is structured as follows: After introducing the data basis and describing the method to test for robustness and to visualize the results (Sections 2 and 3), an application of the method to projected changes of seasonal and extreme precipitation for Germany is presented in Section 4. The results for two ensembles based on the representative concentration pathways (RCPs) 4.5 and 8.5 [14] are compared to the results for an ensemble of simulations based on the emission scenario A1B as defined in the Special Report on Emission Scenarios (SRES [15]). Concluding remarks are given in Section 6.

2. Database

2.1. The Ensemble of Simulations

Owing to their regional focus, climate signal maps are based on data from high-resolution regional climate change simulations. Compared to coarser scale global model simulations, regional simulations
allow for a better representation of mesoscale atmospheric phenomena and of landscape and surface features, such as mountain ranges, land-sea distribution and land cover, which are modifying local wind regimes, temperature patterns, local or regional circulation patterns and precipitation features. A comprehensive review on regional climate models and their applications, including some information on their potentials and limitations, is given by [16]. For this study, we used three different ensembles of regional climate change simulations: one ensemble of 15 simulations of the SRES A1B scenario [15], as provided by the EU FP6 ENSEMBLES project [17], one ensemble of 10 simulations based on the representative concentration pathway (RCP) 4.5 and one ensemble of 11 simulations based on RCP8.5 [14], both provided via the EURO-CORDEX initiative [18,19]. For the RCP8.5 and RCP4.5 ensembles, all data that were available from the Earth System Grid Federation (ESGF) database in August 2014 were included in the analysis. The usage of ENSEMBLES simulations was restricted to those covering the future time period until 2100. The list of simulations is given in Table 1.

Table 1. List of simulations. GCM, global climate model; RCM, regional climate model; RCP, representative concentration pathway.

<table>
<thead>
<tr>
<th>Project/Initiative</th>
<th>Driving Data (GCM)</th>
<th>Regional Model (RCM)</th>
<th>Regional Modeling Group</th>
<th>Scenario/RCP</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENSEMBLES</td>
<td>HadCM3Q3</td>
<td>RCA3.0</td>
<td>SMHI</td>
<td>A1B</td>
</tr>
<tr>
<td>ENSEMBLES</td>
<td>HadCM3Q0</td>
<td>HadRM3Q0</td>
<td>HC</td>
<td>A1B</td>
</tr>
<tr>
<td>ENSEMBLES</td>
<td>HadCM3Q3</td>
<td>HadRM3Q3</td>
<td>HC</td>
<td>A1B</td>
</tr>
<tr>
<td>ENSEMBLES</td>
<td>HadCM3Q16</td>
<td>HadRM3Q16</td>
<td>HC</td>
<td>A1B</td>
</tr>
<tr>
<td>ENSEMBLES</td>
<td>HadCM3Q0</td>
<td>CLM</td>
<td>ETHZ</td>
<td>A1B</td>
</tr>
<tr>
<td>ENSEMBLES</td>
<td>HadCM3Q1-1</td>
<td>RCA3</td>
<td>C4I</td>
<td>A1B</td>
</tr>
<tr>
<td>ENSEMBLES</td>
<td>ECHAM5/MIPO-M</td>
<td>RCA3.0</td>
<td>SMHI</td>
<td>A1B</td>
</tr>
<tr>
<td>ENSEMBLES</td>
<td>ECHAM4/MIPO-M</td>
<td>REMO</td>
<td>MPI-M</td>
<td>A1B</td>
</tr>
<tr>
<td>ENSEMBLES</td>
<td>ECHAM5/MIPO-M</td>
<td>RACMO2</td>
<td>KNMI</td>
<td>A1B</td>
</tr>
<tr>
<td>ENSEMBLES</td>
<td>ECHAM5/MIPO-M</td>
<td>HIRHAM5</td>
<td>DMI</td>
<td>A1B</td>
</tr>
<tr>
<td>ENSEMBLES</td>
<td>ECHAM5/MIPO-M</td>
<td>REGCM3</td>
<td>ICTP</td>
<td>A1B</td>
</tr>
<tr>
<td>ENSEMBLES</td>
<td>BCM</td>
<td>RCA3.0</td>
<td>DMI</td>
<td>A1B</td>
</tr>
<tr>
<td>ENSEMBLES</td>
<td>BCM</td>
<td>HIRHAM5</td>
<td>DMI</td>
<td>A1B</td>
</tr>
<tr>
<td>ENSEMBLES</td>
<td>ARPEGE</td>
<td>HIRHAM5</td>
<td>DMI</td>
<td>A1B</td>
</tr>
<tr>
<td>ENSEMBLES</td>
<td>ARPEGE</td>
<td>RM5.1</td>
<td>CNRM</td>
<td>A1B</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>CNRM-CM5</td>
<td>CCLM4-8-17</td>
<td>CLM community</td>
<td>RCP4,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>CNRM-CM5</td>
<td>RCA4</td>
<td>SMHI</td>
<td>RCP4,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>EC-EARTH</td>
<td>CCLM4-8-17</td>
<td>CLM community</td>
<td>RCP4,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>EC-EARTH</td>
<td>RCA4</td>
<td>SMHI</td>
<td>RCP4,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>EC-EARTH</td>
<td>RACMO22E</td>
<td>KNMI</td>
<td>RCP4,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>EC-EARTH</td>
<td>HIRHAM5</td>
<td>DMI</td>
<td>RCP4,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>CM5A-MR</td>
<td>WRF331F</td>
<td>IPSL</td>
<td>RCP4,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>HadGem2</td>
<td>RCA4</td>
<td>SMHI</td>
<td>RCP4,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>MPI-ESM-LR</td>
<td>CCLM4</td>
<td>CLM community</td>
<td>RCP4,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>MPI-ESM-LR</td>
<td>REMO2009</td>
<td>MPI-M and CS2.0</td>
<td>RCP4,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>CNRM-CM5</td>
<td>CCLM4-8-17</td>
<td>CLM community</td>
<td>RCP8,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>CNRM-CM5</td>
<td>RCA4</td>
<td>SMHI</td>
<td>RCP8,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>EC-EARTH</td>
<td>CCLM4-8-17</td>
<td>CLM community</td>
<td>RCP8,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>EC-EARTH</td>
<td>RCA4</td>
<td>SMHI</td>
<td>RCP8,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>EC-EARTH</td>
<td>RACMO22E</td>
<td>KNMI</td>
<td>RCP8,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>EC-EARTH</td>
<td>HIRHAM5</td>
<td>DMI</td>
<td>RCP8,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>CM5A-MR</td>
<td>RCA4</td>
<td>SMHI</td>
<td>RCP8,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>HadGem2-ES</td>
<td>RCA4</td>
<td>SMHI</td>
<td>RCP8,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>MPI-ESM-LR</td>
<td>CCLM4-8-17</td>
<td>CLM community</td>
<td>RCP8,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>MPI-ESM-LR</td>
<td>REMO2009</td>
<td>MPI-M and CS2.0</td>
<td>RCP8,5</td>
</tr>
<tr>
<td>EURO-CORDEX</td>
<td>MPI-ESM-LR</td>
<td>RCA4</td>
<td>SMHI</td>
<td>RCP8,5</td>
</tr>
</tbody>
</table>
The differences and similarities between the RCP and SRES scenarios are explained, e.g., in the first chapter of the Working Group 1 contributions to the Fifth Assessment Report [20], where Figures 1–15 show the radiative forcing together for the RCP and SRES scenarios. From the three scenarios included in our analysis, RCP4.5 has the lowest radiative forcing, which increases until about 2060 and stays on a constant level of about 4 W/m² thereafter. Contrarily, both RCP8.5 and SRES A1B have increasing radiative forcing until 2100, with A1B showing noticeable lower values than RCP8.5 from about 2050 on. In 2100, RCP8.5 shows a radiative forcing of about 8 W/m² and A1B about 6 W/m².

The ENSEMBLES data are provided on a 25 km × 25 km grid; the EURO-CORDEX simulations have a grid resolution of 0.11° (≈12 km × 12 km). All three ensembles include multiple driving GCMs and multiple RCMs, although the application of specific GCMs and RCMs is owed to the participating modeling institutions, with some models (or model families) being present more often than others. Multiple realizations of one GCM/RCM combination are not present in any of the three ensembles, which is the reason why the individual contribution of internal variability and model uncertainty to the overall spread of the simulation results cannot be quantified.

2.2. Climatic Parameters

In this study, we concentrate on seasonal precipitation (prw, winter; prs, summer) and on the number of days exceeding today’s (1971–2000) 95th percentile of daily winter/summer precipitation (r95pw/r95ps). Seasonal percentiles are calculated from all days of the respective season with precipitation exceeding the 1 mm/day threshold. Summer is defined as June/July/August and winter as December/January/February. The climate changes visualized in the climate signal maps have been calculated as the differences between future 30-year periods (2031–2060, 2041–2070, 2051–2080, 2061–2090) and the reference period 1971–2000. Such 30-year periods are commonly used to represent the climate at the respective time. Ideally, they are short enough to have small trends only within the period and long enough to catch the inter-annual and decadal variability of the climatic parameters. To assess whether a 30-year period is suited to represent the internal variability of the climatic parameter, the temporal variance of 30-year data samples (2031–2060) was calculated and compared to 50-year data samples (2021–2070). Before doing so, both data samples were detrended, assuming a linear trend. This was done for all parameters, seasons, regions and model simulations. Assuming that both data samples (30 years and 50 years) show a good representation of the same statistical population of data, similar variances can be expected from both samples. The coefficient of determination $R^2$ was calculated to evaluate whether this assumption is valid. For prs and prw, the 30-year variances and the 50-year variances are of the same order of magnitude, with $R^2$ values between 0.76 (prs, RCP8.5) and 0.97 (prw, RCP8.5), indicating that the 30-year period is suited to catch the variability of the respective parameters. For r95pw and r95ps, the variances in the 30-year period are slightly larger than in the 50-year period. This is also reflected in smaller $R^2$ values, which lie between 0.49 (r95ps, RCP8.5) and 0.71 (r95pw, RCP8.5). This indicates that a period of 30 years is not entirely sufficient to avoid sampling errors when it comes to extremes, but it is still a reasonable compromise, since trends within 50-year periods become too large compared to the trends under investigation.
3. Climate Signal Map Method

As mentioned in Section 1, reduced complexity also means reduced information content. This is achieved by the following measures: (i) only one direction of changes is considered (increase or decrease); in regions where the changes have the opposite sign, no information is displayed; (ii) the changes are graduated into very coarse categories and displayed using only three colors: green for small changes, orange for medium changes and red for large changes; (iii) the projected changes can be displayed using application-oriented regions (see Section 3.1); (iv) in regions where the climate change signal is classified as not robust (see Section 3.2), no information is displayed (these regions are displayed in gray); and (v) in regions where the climate change signal is classified as robust, ensemble median changes are displayed; no information on the variability around this median change is shown.

3.1. Regional Mapping

Climate signal maps are intended to be used by non-climate modelers. Therefore, the information must be displayed in a familiar context. For decision makers from public services, this can be administrative units, such as countries, states or rural districts. For ecologists, physical regions and landscapes are more convenient; economists might think in spatial units as defined by the NUTS (“Nomenclature des unités territoriales statistiques”) system [21]. The method can of course also be applied to the model data on their native model grids (in the case that the simulations are done with similar horizontal grid resolutions) or on a common grid to which all simulation results are mapped. Here, we use the German rural districts to illustrate the method and results (see Section 4).

3.2. Robustness Tests

Climate signal maps allow for statements on one single direction of change, which has to be specified in advance. Robustness is thus also tested only for this distinct sign of the change. We define robustness as a combination of model agreement and the significance of the individual model projections. For a given sub-domain, both tests presented in the following sections must be passed in order for the simulated median change signal to be considered as robust.

3.2.1. Test for Agreement

The first measure for determining the robustness of climate change information given by an ensemble of climate change simulations is the agreement of the simulations in terms of the direction of the changes. This is done in the same way as was presented in the 4th and 5th IPCC Assessment Reports and as it is described in the IPCC AR5 guidance note on consistent treatment of uncertainties [22]. The basic idea of this test is that the climate change signal derived from an ensemble of climate change simulations is more robust if the majority of the simulations agree on the sign of the changes than if half of the simulations project positive changes and the other half project negative changes. Depending on the application and on the data basis, the strength of this measure of robustness can be chosen via the percentage of simulations that have to agree in sign. For the examples shown in Section 4, an agreement of at least 66% of the simulations is required to pass the test. This relatively low percentage of simulations that have to agree
on the sign of the changes is mainly owed to the small number of simulations building the ensembles. For small ensemble sizes, requesting high percentages of agreement causes test failure, even if only a few simulations deviate considerably from the rest of the data.

3.2.2. Test for Significance

The second measure of robustness is the statistical significance of the simulated climate changes. The U-test (Mann–Whitney–Wilcoxon test, [23–25]) is applied to investigate whether for a climatic parameter (e.g., the number of days per year with precipitation exceeding the 95th percentile of daily precipitation) the future distribution differs from the distribution for present day climate, that is the climate change signal is significant. The basic idea of this test is as follows: the data samples for present and for future climate are combined to assess whether the ranks of the data samples are shifted towards higher or lower ranks for future climate (resulting in a positive or a negative climate change signal) or whether the ranks are equally distributed among both samples (resulting in no climate change signal). In the case of climate signal maps, the 30 years contributing to the present-day climate period and those constituting the 30-year future climate period are used as samples for the Mann–Whitney–Wilcoxon test statistics. For the following reasons, the test is especially suited for this application:

1. by using the ranks, the test attains robustness versus possible distributional shapes of the sample data, which also makes it applicable for right-skewed distributions, such as for precipitation
2. it is intuitively comprehensible and widely used in many scientific disciplines [26,27].

One condition for the U-test is the independence of the data of one sample. For this reason, for each parameter, autocorrelation was calculated for each 30-year data sample (i.e., for n simulations, t time periods and r regions, resulting in a n × r × t time series). Apart from the zero time lag, which gives a perfect autocorrelation of one, the autocorrelation values are relatively small (in the range of −0.3–0.3), regardless of the time lag. Similar results are obtained for all parameters, regions and seasons. When introducing new parameters into the climate signal maps, the compliance of this condition has to be tested again.

The U-test was carried out for each model simulation separately. To make a statement on the robustness of the whole simulation ensemble, it is necessary that a given fraction of the ensemble members shows significant results in a region of interest. For the examples shown in Section 4, the condition to pass the significance test is that at least 66% of the simulations show significant changes with a significance level of 0.85. Again, depending on the application and on the data basis, the significance level, as well as the percentage of simulations that have to pass the test can be adjusted.

4. Results

To illustrate the method of the climate signal maps, we show two types of figures. Firstly, time series of the projected changes are shown, which are based on 30-year running mean anomalies with respect to the reference period of 1971–2000 for each ensemble member. The data have been spatially averaged for the area of Germany. The figures show the envelope of the projections of the entire ensemble, i.e., the lines denote the highest and the lowest changes. Note that values for different time steps may come
from different models. The shaded areas represent the bandwidth given by the ensemble of simulations. Secondly, climate signal maps are shown for four different 30-year periods (2031–2060, 2041–2070, 2051–2080, 2061–2090), again with respect to the reference period of 1971–2000. Ensemble median changes are shown in the top panels, with low changes colored in green, medium changes in orange and larger changes in red. Ensemble median changes are used, because they are less influenced by single outliers than ensemble mean changes. For the lower panels, the robustness tests as described in Section 3.2 were applied. Regions that failed in at least one of the tests are shown in gray.

The projected changes in winter precipitation for the EURO-CORDEX RCP4.5 and RCP8.5 simulations and for the ENSEMBLES A1B simulations are shown in Figure 1. The EURO-CORDEX simulations mostly project increasing winter precipitation, which gets more pronounced for the simulations based on RCP8.5 after 2040. However, the bandwidth of changes of the RCP4.5 simulations includes the zero-changes line for the entire period; for the RCP8.5 simulations, this is the case until about 2040. The bandwidth spanned by the ENSEMBLES simulations is much broader, already for the early decades from 2020.

![Figure 1](image_url)

**Figure 1.** Projected change in winter precipitation (%). The figure shows the envelope of 30-year running mean anomalies with respect to the reference period of 1971–2000 for the ENSEMBLES A1B simulations (light gray area and dotted lines), for the EURO-CORDEX RCP8.5 simulations (medium gray area and dashed lines) and for the EURO-CORDEX RCP4.5 simulations (dark gray area and solid lines). The temporal means are mapped on the central year of the respective 30-year period.

For the same data, climate signal maps are shown for the RCP4.5 (Figure 2), RCP8.5 (Figure 3) and A1B (Figure 4) simulations. Interestingly, for the RCP4.5-based simulations, there are regions in southwestern Germany where the projected changes in winter precipitation are robust already for the near future time periods (2031–2060, 2041–2070) and stay robust until the end of the 21st century. For the RCP8.5-based simulations, regions with robust changes do not emerge prior to the 2041–2070 period, but cover larger parts of the country and show stronger changes.
**Figure 2.** Increase of winter precipitation (%) for (from left to right) 2031–2060, 2041–2070, 2051–2080 and 2061–2090 compared to the reference period 1971–2000. Color coded: median of 10 RCM simulations from the EURO-CORDEX RCP4.5 simulations. For the top panels, no test was applied to the data. In the lower panels, regions that failed at least one of the two robustness tests are grayed out.

**Figure 3.** Increase of winter precipitation (%) for (from left to right) 2031–2060, 2041–2070, 2051–2080 and 2061–2090 compared to the reference period 1971–2000. Color coded: median of 11 RCM simulations from the EURO-CORDEX RCP8.5 simulations. For the top panels, no test was applied to the data. In the lower panels, regions that failed at least one of the two robustness tests are grayed out.
The ENSEMBLES A1B simulations (Figure 4) differ substantially from the EURO-CORDEX simulations (Figures 2 and 3). Firstly, the regional patterns of winter precipitation increase differ. For the EURO-CORDEX simulations, the largest increase is projected for southern Germany for both the RCP4.5 and RCP8.5 ensemble. From the ENSEMBLES simulations, the strongest increase emerges for northeast Germany. This effect can be related to differences in the simulated changes in large-scale circulation patterns between the CMIP3 models [28], which are the basis for the regional ENSEMBLES simulations and the CMIP5 models, building the basis for the EURO-CORDEX simulations [29]. Secondly, the ENSEMBLES A1B projections for increasing winter precipitation are identified as robust already for the early time periods for most of Germany. This is unexpected considering the large bandwidth of the projections (Figure 1). However, the bandwidth might be misleading in this case, as the negative changes are projected by two simulations only (both driven by the same GCM). The test for agreement was passed, as two simulations out of an ensemble of 15 simulations are less than the requested 34% deviant changes, which would cause a test failure.

The projected changes in summer precipitation are shown in Figures 5–8. From the time series (Figure 5), no clear trend of domain-averaged summer precipitation change (neither an increase, nor decrease) can be derived. Instead, all three ensembles enclose the zero-changes line in their bandwidth, with highest increases projected by the ensemble of RCP4.5 simulations towards the end of the 21st century. As for winter precipitation, the ENSEMBLES ensemble again gives a broader bandwidth with stronger decreases towards the end of the 21st century.
**Figure 5.** Projected change in summer precipitation (%). The figure shows the envelope of 30-year running mean anomalies with respect to the reference period of 1971–2000 for the ENSEMBLES A1B simulations (light gray area and dotted lines), for the EURO-CORDEX RCP8.5 simulations (medium gray area and dashed lines) and for the EURO-CORDEX RCP4.5 simulations (dark gray area and solid lines). The temporal means are mapped on the central year of the respective 30-year period.

**Figure 6.** Decrease of summer precipitation (%) for (from left to right) 2031–2060, 2041–2070, 2051–2080 and 2061–2090 compared to the reference period 1971–2000. Color coded: median of 10 RCM simulations from the EURO-CORDEX RCP4.5 simulations. For the top panels, no test was applied to the data. In the lower panels, regions that failed at least one of the two robustness tests are grayed out.
Figure 7. Decrease of summer precipitation (%) for (from left to right) 2031–2060, 2041–2070, 2051–2080 and 2061–2090 compared to the reference period 1971–2000. Color coded: median of 11 RCM simulations from the EURO-CORDEX RCP8.5 simulations. For the top panels, no test was applied to the data. In the lower panels, regions that failed at least one of the two robustness tests are grayed out.

Figure 8. Decrease of summer precipitation (%) for (from left to right) 2031–2060, 2041–2070, 2051–2080 and 2061–2090 compared to the reference period 1971–2000. Color coded: median of 15 RCM simulations from the ENSEMBLES A1B simulations. For the top panels, no test was applied to the data. In the lower panels, regions that failed at least one of the two robustness tests are grayed out.
Here, the climate signal map method is used to check for the existence of regions where a robust decrease of the summer precipitation can be derived from the data. This is shown in Figures 6–8. For RCP4.5 (Figure 6), the median decrease of summer precipitation is weak and basically limited to southern Germany until about 2075 and to western Germany thereafter. As could be expected from Figure 5, the climate signal map method does not indicate any robust decrease in the projected summer precipitation for the RCP4.5 simulations.

For RCP8.5, however, the climate signal maps indicate regions with robust projections of decreasing summer precipitation. Figure 7 shows temporally expanding and intensifying regions in southwestern Germany where summer precipitation is projected to decrease and which start to become robust towards the end of the 21st century. The ENSEMBLES ensemble projects decreases in summer precipitation that are relatively similar to those projected by the EURO-CORDEX RCP8.5 ensemble in terms of location and intensity of the strongest changes.

The comparatively low radiative forcing of RCP4.5 translates into low and insignificant decreases of summer precipitation, whereas the stronger radiative forcing of RCP8.5 and A1B leads to regions with stronger and robust decreases of summer precipitation in southeastern Germany.

For the number of days exceeding the 95th percentile of today’s daily winter precipitation (r95pw), all three ensembles show a positive trend with the largest increases projected for the RCP8.5 ensemble (Figure 9). This is also reflected in the climate signal maps (Figures 10–12), which show strong and mostly robust increases for all three ensembles towards the end of the 21st century. The RCP4.5 ensemble projects the lowest increase and lowest robustness of the changes with robust increases for significant parts of Germany not prior to 2065. In contrast, both the RCP8.5 and the ENSEMBLES ensemble project robust changes of r95pw for large parts of Germany already for the period centered at 2055.

![Figure 9. Change in the number of winter days exceeding today’s 95th percentile of daily winter precipitation (%) for the EURO-CORDEX RCP4.5 and RCP8.5 simulations. The figure shows the envelope of 30-year running mean anomalies with respect to the reference period of 1971–2000 for the ENSEMBLES A1B simulations (light gray area and dotted lines), for the EURO-CORDEX RCP8.5 simulations (medium gray area and dashed lines) and for the EURO-CORDEX RCP4.5 simulations (dark gray area and solid lines). The temporal means are mapped on the central year of the respective 30-year period.](image-url)
Figure 10. Increase of the number of winter days exceeding today’s 95th percentile of daily winter precipitation (%) for (from left to right) 2031–2060, 2041–2070, 2051–2080 and 2061–2090 compared to the reference period 1971–2000. Color coded: median of 10 RCM simulations from the EURO-CORDEX RCP4.5 simulations. For the top panels, no test was applied to the data. In the lower panels, regions that failed at least one of the two robustness tests are grayed out.

Figure 11. Increase of the number of winter days exceeding today’s 95th percentile of daily winter precipitation (%) for (from left to right) 2031–2060, 2041–2070, 2051–2080 and 2061–2090 compared to the reference period 1971–2000. Color coded: median of 11 RCM simulations from the EURO-CORDEX RCP8.5 simulations. For the top panels, no test was applied to the data. In the lower panels, regions that failed at least one of the two robustness tests are grayed out.
**Figure 12.** Increase of the number of winter days exceeding today’s 95th percentile of daily winter precipitation (%) for (from left to right) 2031–2060, 2041–2070, 2051–2080 and 2061–2090 compared to the reference period 1971–2000. Color coded: median of 15 RCM simulations from the ENSEMBLES A1B simulations. For the top panels, no test was applied to the data. In the lower panels, regions that failed at least one of the two robustness tests are grayed out.

**Figure 13.** Change in the number of summer days exceeding today’s 95th percentile of daily summer precipitation (%) for the EURO-CORDEX RCP4.5 and RCP8.5 simulations. The figure shows the envelope of 30-year running mean anomalies with respect to the reference period of 1971–2000 for the ENSEMBLES A1B simulations (light gray area and dotted lines), for the EURO-CORDEX RCP8.5 simulations (medium gray area and dashed lines) and for the EURO-CORDEX RCP4.5 simulations (dark gray area and solid lines). The temporal means are mapped on the central year of the respective 30-year period.
The changes in the number of summer days exceeding the 95th percentile of today’s daily summer precipitation (r95ps) are shown in Figure 13. For RCP4.5 and RCP8.5, the trend to rising numbers of days exceeding the 95th percentile is visible, although for both ensembles, the zero-changes line is included in the bandwidth. The climate signal maps show positive median changes for both ensembles, which are however not robust for either period. The A1B ensemble also shows only very few robust regions of increasing r95ps. However, in contrast to the two RCP ensembles, the climate signal maps based on the A1B ensemble indicate regions in western Germany, where r95ps either decreases or stays constant (not shown). This is also reflected in the stronger negative branch of the bandwidth of the ENSEMBLES A1B ensemble compared to the two RCP ensembles shown in Figure 13.

5. Discussion

Methods to assess the robustness of climate change information are always based on some assumptions and prerequisites. Therefore, one should also think about the robustness of the robustness. Here, we shortly discuss some major points that can easily change the outcome of analyses like ours. No matter how sophisticated the method to determine the robustness of projected climate changes is, the result always depends on the composition of the underlying ensemble of simulations. Despite the efforts undertaken to structure the matrix of simulations used in initiatives, such as CMIP and CORDEX, or projects, like ENSEMBLES, the ensembles of simulations are basically ensembles of opportunity, which are composed rather arbitrarily and affected by the interdependencies among models. For instance, the CMIP5 ensemble comprises nearly identical models from some institutions that were found to artificially increase the robustness [11]. There are approaches to circumvent the problems associated with model interdependency or over-representation of certain models in an ensemble, like data reconstruction (e.g., [30,31]) or model selection (e.g., [32]). However, there is no commonly-accepted method available yet, and we therefore preferred to leave the ensemble as it is and to keep as many simulations as possible. Contrarily, as seen in Section 4 for the decrease in summer precipitation from the ENSEMBLES ensemble, some simulations produce a very different climate change signal than the others. Either this is due to poor model quality (in this case, they can reduce robustness by mistake) or they are superior to others (e.g., being the only simulations done by models that represent processes that strongly influence climate sensitivity), and these models are dramatically underrepresented in the ensemble, so that they might be attributed by mistake as “wrong outliers”. The climate signal map method is not able to decide if the underlying ensemble of simulations is well chosen, and the interpretation always has to be based on the facts that the ensemble might not represent the full bandwidth of possible future climate evolutions and that the ensemble might be dominated by single models (or model families).

Especially in the upcoming decades, the climate change signal might be relatively small, making it difficult to differentiate between internal variability and a climate change signal. If internal variability is not correctly represented by the ensemble (too few or no independent models, too few realizations of individual GCM-RCM combinations, simplified or missing processes in the models, etc.), this can lead to both wrongly-attributed “robust” and “non-robust” regions. The first can happen if the variability represented by the ensemble is too small, the latter if the internal variability given by the ensemble of simulations is too large. An example can be seen in Figures 2 and 3 for the near future time period of
Although the underlying simulations ensemble is very similar in terms of models and the size of the ensemble and RCP4.5 and RCP8.5 greenhouse gas concentrations are also very similar at this point of time, the climate signal maps for 2031–2060 look very different. For the RCP4.5 ensemble, large parts in southwestern Germany are determined to show robust climate changes, while the RCP8.5 ensemble results in very few regions with robust changes.

Furthermore, the grid resolution of the ensemble of simulations has to be suited for the size of the regions and for the chosen climatic parameter. For small regions as, e.g., the German administrative units used in this study, high-resolution climate model data have to be used; the same accounts for regionally very heterogeneous parameters, such as precipitation extremes, which might not be represented by model simulations on coarser grid resolutions.

The spatial aggregation from grid boxes to regions has some consequences. Firstly, the regions might be of unequal size. By averaging the parameters in each region, a very large region might not be comparable to a very small region, especially for very heterogeneous parameters, because the extremes are spatially smoothed. Secondly, some of the regions can be too small to be represented properly by the model grid size. For spatially-heterogeneous parameters, it is advisable to not consider the values given for one single grid box, but to take at least the averages over 4–9 grid boxes [33]. Therefore, for very small regions, the neighboring regions should also be taken into account for the interpretation of the results. Furthermore, the unavoidable loss of spatial information by aggregating the data from grid boxes to regions must be considered. Especially for locally-heterogeneous parameters in combination with large regions, local information gets lost. Despite these possible drawbacks, we apply a regional mapping for the climate signal maps for the sake of usability. A stakeholder from a district administration will easily understand what a mean reduction in winter precipitation of, say, 10% means for his or her district, but could have problems identifying and extracting information for his region of interest from figures based on data on typical climate model grids.

The climate change signals are calculated based on 30-year periods, which are supposed to represent a climate state at that time. This is problematic (i) if the climatic parameter shows a very strong trend during this 30-year period, because in this case, the period represents rather a transient climate evolution than a stationary state, or (ii) if the parameter is extremely heterogeneous in terms of its temporal distribution, being either a very rare event or having a large variability, making it difficult to capture the statistical distribution within a 30-year period. In both cases, more sophisticated methods of trend analysis and extreme value theory should be applied to test for climate change signals and their robustness.

Finally, the climate signal map method allows one to adjust some parameters, such as the significance level of the significance test or the fraction of simulations that have to agree on the direction of changes or that have to pass the significance test. These parameters can be chosen according to the application. Too restrictive requirements (e.g., 100% model agreement combined with 99% significance) will yield close to no robust information. Hence, the user has to find a compromise. Climate signal maps help with finding the compromise by allowing one to adjust those parameters. To illustrate the sensitivity of climate signal maps to those adjustable parameters, the percentage of regions passing the two individual tests was calculated with different parameter settings. The results are shown in Table 2, where the percentage of regions passing the individual test is given for winter precipitation (prw, RCP4.5 and
RCP8.5) and for the number of winter days exceeding the 95th percentile of winter precipitation (r95pw, RCP4.5 and RCP8.5). The test for agreement was performed with two settings: 66% and 90% required simulation agreement on the sign of the changes ($P_{ag}$). The significance test was performed in four settings, altering the required percentage of simulations that have to show significant changes ($P_{sig}$) and the required significance level ($L_{sig}$). For the climate signal maps, a region is determined to show robust climate changes only if both tests are passed. The future time period used for this sensitivity study was 2061–2090. The results (see Table 2) show that the large model agreement of the RCP8.5 simulations leads to only small losses of robust regions when increasing $P_{ag}$ from 66% to 90%. For the RCP4.5 simulations, model agreement is smaller, resulting in stronger decreases in the number of robust regions. In the case of the test for significance, increasing the percentage of regions that have to show significant changes from 66% to 90% has in all cases a stronger influence on the number of robust regions than increasing the significance level from 0.85–0.90. It can also be seen that, independent of the parameter setting, the test for significance is passed by less regions than the test for agreement, being thus the more critical test for the shown examples. In general, the sensitivity to the setting of the adjustable parameters depends strongly on the heterogeneity of the ensemble of simulations, on the variability of the climatic parameter and on the strength of the climate change signal.

Table 2. The sensitivity of the robustness tests to the setting of the adjustable parameters for projected changes (2061–2090 wrt.1971–2000) of the 95th percentile of daily winter precipitation (r95pw) (RCP8.5 and RCP4.5) and prw (winter precipitation) (RCP8.5 and RCP4.5). The numbers denote the percentage of regions that are defined as robust under the test conditions given in the leftmost table column with $P_{ag}$: percentage of simulations that have to agree on the sign of the projected changes; $P_{sig}$: percentage of simulations that have to show significant changes.; and $L_{sig}$: the significance level.

<table>
<thead>
<tr>
<th></th>
<th>r95pw RCP8.5</th>
<th>r95pw RCP4.5</th>
<th>prw RCP8.5</th>
<th>prw RCP4.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{ag} = 66%$</td>
<td>99</td>
<td>100</td>
<td>99</td>
<td>100</td>
</tr>
<tr>
<td>$P_{ag} = 90%$</td>
<td>96</td>
<td>88</td>
<td>97</td>
<td>69</td>
</tr>
<tr>
<td>$P_{sig} = 66%$, $L_{sig} = 0.85$</td>
<td>95</td>
<td>70</td>
<td>84</td>
<td>57</td>
</tr>
<tr>
<td>$P_{sig} = 66%$, $L_{sig} = 0.90$</td>
<td>90</td>
<td>40</td>
<td>67</td>
<td>30</td>
</tr>
<tr>
<td>$P_{sig} = 90%$, $L_{sig} = 0.85$</td>
<td>64</td>
<td>10</td>
<td>52</td>
<td>6</td>
</tr>
<tr>
<td>$P_{sig} = 90%$, $L_{sig} = 0.90$</td>
<td>45</td>
<td>6</td>
<td>35</td>
<td>5</td>
</tr>
</tbody>
</table>

6. Conclusions

Climate signal maps identify regions with robust climate changes from an ensemble of climate change simulations. Here, robustness is defined as a combination of model agreement and significance of the individual model projections. Climate signal maps do not show all information available from the model ensemble, but give a condensed view in order to be useful for non-climate scientists who have to assess climate change impact during the course of their work. By not showing ensemble median changes in regions where the projected changes are determined to be not robust (gray regions), the climate signal maps show that it is not possible to deliver robust information on expected climate changes for every
parameter, time period and region. They can therefore be used to illustrate the fact that climate change impact assessment and adaptation to climate change always has to cope with some unknowns. The selection of the ensemble of simulations building the basis of the climate signal maps has to be done with care, taking both underlying scenario assumptions and the imperfectness of the ensemble in terms of sampling errors, model interdependencies and the coverage of the full bandwidth of possibilities into account.

In this work, we presented the results for three different ensembles of regional climate change simulations. Although the models used and the scenario assumptions differ (RCP4.5 vs. RCP8.5 vs. A1B), some similarities in the projections of future seasonal and extreme precipitation can be seen across the different scenarios and models. All three ensembles project increasing winter precipitation. However, the regional patterns differ between the RCP ensembles (strongest increase in the southern parts of Germany) and the A1B ensemble (strongest increase in northeast Germany). Furthermore, the strength of the projected changes (weakest for RCP4.5 and strongest for A1B), as well as the robustness (some robust regions from the beginning, which become only slightly larger toward the end of the century for RCP4.5, increasing robustness towards the end of the century for RCP8.5 and robust changes for all time periods for most of Germany for A1B) differ between the different ensembles of simulations. A robust decrease of summer precipitation can be derived only for small regions in southwestern Germany from the A1B ensemble and the RCP8.5 simulations. For RCP4.5, large parts of Germany do not show any decreases in summer precipitation at all; those with negative changes cannot be attributed as robust. Regarding winter precipitation extremes, all three ensembles show a positive and mostly robust trend of the number of winter days exceeding the 95th percentile of today’s daily winter precipitation with the largest increases projected by the RCP8.5 ensemble. Contrarily, the projected increase of the number of summer days exceeding the 95th percentile of today’s daily summer precipitation is relatively small and mostly not robust for all ensembles and time periods.

Our findings for the robustness of projected changes of seasonal precipitation are consistent with previous studies (e.g., [17,19,34]), which have shown the strong seasonality of the projected changes with a trend towards dryer summers and wetter winters and a strong north-south gradient, with drying in the south and no changes or even increasing precipitation in the north of Europe. The transition zone between the northern European regions, mostly confronted with increasing winter precipitation, and the southern European regions, facing the problem of decreasing summer precipitation, is located over central Europe.

The differences of the pattern of winter precipitation changes between the RCP ensembles and the ENSEMBLES A1B ensemble, possibly influenced by differences in the projected large-scale circulation from the CMIP3 and CMIP5 models, is worth being addressed in future research.

An interesting feature is the consistently larger bandwidth in the ENSEMBLES A1B ensemble compared to the two RCP ensembles, which is probably caused by the composition of the global and regional models used for the simulations. However, the larger bandwidth does not necessarily mean lower robustness, as it can be caused by a few outliers. For such cases, the method of climate signal maps proved to be a useful tool to identify regions with robust signals despite the large bandwidth of the ensemble.
In this work, we showed the application of the climate signal map method to projected future precipitation. The method is also suited for other climatic parameters, provided that the preconditions of the statistical test are matched. If needed, additional tests for robustness can easily be implemented. As an example, the assessment of the robustness of the projected climate changes with respect to the internal natural variability of the climatic parameter could be incorporated into the method. This would, however, require a reliable quantification of the internal variability, which requires either a large and well-defined ensemble of simulations or some advanced methods to derive internal variability from the simulation ensembles available (e.g., [3]).

Acknowledgments

The ENSEMBLES data used in this work were funded by the EU FP6 Integrated Project ENSEMBLES (Contract Number 505539), whose support is gratefully acknowledged. We acknowledge the World Climate Research Programme’s Working Group on Regional Climate and the Working Group on Coupled Modeling, the former coordinating body of CORDEX and the panel responsible for CMIP5. We also thank the climate modeling groups (listed in Table 1 of this paper) for producing and making available their model output. We also acknowledge the Earth System Grid Federation infrastructure, an international effort led by the U.S. Department of Energy’s Program for Climate Model Diagnosis and Intercomparison, the European Network for Earth System Modeling and other partners in the Global Organisation for Earth System Science Portals (GO-ESSP). The German administrative boundaries were provided by the German Federal Agency for Cartography and Geodesy. We gratefully acknowledge the helpful comments from the three anonymous reviewers that improved the paper significantly.

Author Contributions

Susanne Pfeifer did the data analysis and visualization and wrote the manuscript. All authors contributed to the discussion and development of the climate signal map method. Manfred Mudelsee was in charge of assessing the validity of the statistical methods used in the analysis.

Conflicts of Interest

The authors declare no conflict of interest.

References


17. Van der Linden, P.; Mitchell, J. *ENSEMBLES: Climate Change and Its Impacts: Summary of Research And Results from the ENSEMBLES Project*; Met Office Hadley Centre: Exeter, UK, 2009; pp. 1–160.


© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).