Daily Temperature Changes and Variability in ENSEMBLES Regional Models Predictions: Evaluation and Intercomparison for the Ebro Valley (NE Iberia)

A. El Kenawy1,4*, J.I. López-Moreno2, M. F. McCabe1, N.A. Brunsell3 & S.M. Vicente-Serrano2

1. Hydrological Modeling & Earth Observation Group, WDRC, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia.
2. Instituto Pirenaico de Ecología- CSIC, Campus de Aula Dei, P.O. Box 13034, Zaragoza, 50059, Spain.
3. Department of Geography - Atmospheric Sciences Program, University of Kansas, Lawrence, KS 66045-7613, USA.
4. Department of Geography, Mansoura University, Mansoura, Egypt.

* kenawy@mans.edu.eg

We employ a suite of regional climate models (RCMs) to assess future changes in summer (JJA) maximum temperature (Tmax) over the Ebro basin, the largest hydrological division in the Iberian Peninsula. Under the A1B emission scenario, future changes in both mean values and their corresponding time varying percentiles were examined by comparing the control period (1971-2000) with two future time slices: 2021-2050 and 2071-2100. Here, the rationale is to assess how lower/upper tails of temperature distributions will change in the future and whether these changes will be consistent with those of the mean. The model validation results demonstrate significant differences among the models in terms of their capability to representing the statistical characteristics (e.g., mean, skewness and asymmetry) of the observed climate. The results also indicate that the current substantial warming observed in the Ebro basin is expected to continue during the 21st century, with more intense warming occurring at higher altitudes and in areas with greater distance from coastlines. All models suggest that the region will experience significant positive changes in both the cold and warm tails of temperature distributions. However, the results emphasize that future changes in the lower and upper tails of the summer Tmax distribution may not follow the same warming rate as the mean condition. In particular, the projected changes in the warm tail of the summer Tmax are shown to be significantly larger than changes in both mean values and the cold tail, especially at the end of the 21st century. The finding suggests that much of the changes in the summer Tmax percentiles will be driven by a shift in the entire distribution of temperature rather than only changes in the central tendency. Better understanding the possible implications of future climate systems provides
information useful for vulnerability assessments and the development of local adaptation strategies for multi-disciplinary investigations.

**Key words:** Climate models, ENSEMBLES, Trends, Temperature percentiles, Maximum temperature, The Ebro valley, Iberia.

**1. Introduction**

Climate change has become a major focus of attention in the scientific community over the last few decades, due in part to the potential impacts on both natural systems (Burlando and Rosso, 1991; Wood et al., 2004; Ryan et al., 2008) and human environments (Patz et al., 2005). However, although climate change is a global phenomenon, the potential impacts are not homogenous over space (Trenberth, 2011). In particular, the magnitude and impacts of climate change can vary significantly from one region to another. Characterizing changes in climatic patterns and their possible future changes at finer spatial scales can assist in providing needed guidance and advice to local policy makers for adaption and mitigation strategies. In addition, assessing climate change and variability on local and regional scales is critical for improved management and understanding of ecosystem dynamics and human settlements.

Recently, a number of European projects have designed experiments to downscale regional climate projections from Global Climate Models (GCMs), using either dynamical or statistical approaches. Among these projects are PRUDENCE (Christensen et al., 2007), STARDEX (Goodess et al., 2005), ENSEMBLES (van der Linden and Mitchell, 2009), MISTRA-SWECIA (Kjellström, 2008) and EURO-CORDEX (Kotlarski, 2014). These projects allow not only for an improvement in the spatial resolution of climate simulations over Europe, but also for incorporating additional physical processes into model simulations (e.g., cloud formation, aerosol influences, radiation balance, atmospheric fronts and boundary layer processes). Improvement in the capability of these models makes them a promising tool for assessing the potential impacts of future climate change on regional scales across Europe. For this reason, these climate simulations have been investigated quite thoroughly in numerous climate studies over different regions (see for example, Olsson et al., 2009; Van der Linden and Mitchell, 2009; Lorenz and Jacob, 2010;
In response to recent improvements in the capability of numerical models for predicting future climate, a number of studies have employed these high resolution climate simulations to argue that increased heat stress risk over Europe may be caused by elevated greenhouse gas concentrations (Fischer and Schar, 2009; Barriopedro et al., 2011). For example, Parry (2000) studied the future climatic trend in Europe and found that the mean temperature might increase by 2–6.3°C at the end of this century, with the largest warming occurring over Southern Europe and the Mediterranean regions. In the same context, Schar et al. (2004) indicated that the European summers at the end of the current century will be as warm as the extraordinary 2003 summer. They estimated an increase in summer mean temperature of 10 to 15%, with a tendency for increasing temperature variability. This finding was also supported by Della-Marta et al. (2007), who found that the climate of Western Europe has become more extreme, suggesting that the increase in variance of future summer temperature has already been evident over the last 126 years. Fischer and Schar (2009) also found that changes in daily summer extreme temperatures are likely to be accompanied by changes in both the mean and the variance.

In this work, we provide a detailed assessment of future projections of summer maximum temperature ($T_{\text{max}}$) over the Ebro valley in northeast Iberia, using a suite of regional climate models (RCMs) from the European ENSEMBLES project. The Ebro basin is the largest in the Iberian Peninsula, with a total area of nearly 85,000 km² and is characterized by complex topography and climate, with large regional differences (Cuadrat, 1999). The water of the Ebro is the main supply for irrigation and other agricultural activity, domestic supply, electricity production and industrial development. A comprehensive assessment of temperature change projections at fine resolution over the Ebro catchment is lacking, although numerous studies have placed an emphasis on assessing future temperature projections in other Spanish regions (e.g., Vargas et al., 2008; Lopez Moreno et al., 2008; Brands et al., 2011; Gimeno et al., 2011; González-Aparicio and Hidalgo, 2012; Jerez et al., 2013; Domínguez et al., 2013). To date, the majority of studies in the region were focused primarily on an assessment of future changes in precipitation (e.g., Palatella et al., 2010; García-Garizábal et al., 2014). An exception is the study of Lopez-Moreno et al. (2014),
who analyzed observed and future projections of a set of temperature extremes in the Ebro valley during an extended winter (DJFM). Also, Domínguez et al. (2013) analyzed the capacity of five RCM simulations to represent warm days over eleven basins across Spain, including the Ebro.

In this study, we endeavor to provide a detailed assessment of future temperature changes over the Ebro catchment using simulated data from a set of high resolution RCMs. The focus will be on the summer (JJA) season, as summers in the Ebro valley show the highest temperature changes over the last five decades (0.3°C/decade), with a particularly strong rise in the last decade (El Kenawy et al., 2012). A prolonged warming of summer temperatures in the future may intensify the hydrological cycle in the basin, resulting in higher evaporation losses, higher irrigation water requirements, and an overall increase in water resources demands for domestic, agricultural and industrial use (El Kenawy et al., 2013; Vicente-Serrano et al., 2014a,b). In addition, anomalous heat events in future decades may inflict large damages and losses to the region, including human fatalities, forest fires, water deficits and crop losses.

The main objectives of this work are: (1) to evaluate the capability of a suite of nine RCMs to reproduce near present summer \( T_{\text{max}} \) variations in the Ebro valley, (2) to assess projected changes in summer mean \( T_{\text{max}} \), as forced by the A1B emission scenario for the 21st century, and (3) to explore future changes in summer heat events. The latter objective is undertaken by comparing changes in the lowest (pc1, pc5 and pc10) and highest (pc90, pc95 and pc99) percentiles of the summer \( T_{\text{max}} \) distribution with those of the mean state. This assessment is important to determine whether projected changes in temperature percentiles will be consistent with simulated changes in the corresponding mean climate. Indeed, changes in the warm and cold tails of temperature distribution are interesting, particularly for regional climate impact and assessment studies. Over Iberia, little research has explicitly indicated the extent to which simulated changes in mean temperature can be related to its corresponding probability distribution, particularly on a more localized scale.

2. Data and Methodology

2.1. Study area
The Ebro valley is one of the main physiographic units in the Iberian Peninsula, representing approximately 17% of Spain. The altitude varies from below 200 m to above 2500 m (Figure 1). The Ebro extends inland along a northwest-southeast axis in a semi-enclosed basin, surrounded by mountain belts including the Pyrenees (north), the Cantabrian belt (northwest), the Catalan chain (east) and the Iberian system (south and southwest). This geographical domain is unique from a climatic perspective, as it includes the northern most semi-arid region in the European continent. Furthermore, with the latitudinal location, the dynamics of the mid latitude circulation and the sub-tropics are thoroughly linked and compete against each other. The climate of the domain is influenced mainly by large-scale atmospheric configurations, originating from the north Atlantic and the Mediterranean Sea and including, among others, the North Atlantic sub-tropical high and the Atlantic sea surface temperature (SST). However, the impacts of these large-scale atmospheric circulations are sometimes modulated by local terrain (Vicente-Serrano et al., 2009). The study domain is characterized by a marked temperature gradient during the summer, extending from south to north. The catchment average T\textsubscript{max} over the period 1960-2006 is 27.1°C, but the long-term average T\textsubscript{max} is approximately 36°C in the lower elevations in the central parts of the basin. Daily maximum temperatures at these sites sometimes exceed 40°C, marking them as some of the hottest places in Europe.

2.2. Observed temperature data

In this research, we have used a complete and homogenous daily temperature data set derived from 63 meteorological stations covering the Ebro basin from 1971 to 2000 to validate the simulated data (see details in El Kenawy et al., 2013). In comparison with other data sets, such as E-OBS ECA [Hofstra et al., 2009], EMULATE [Moberg et al., 2006], SDATS [Brunet et al., 2006] and Spain02 [Herrera et al., 2012], the Ebro basin is well sampled by a relatively dense network of observational stations (N=63). This network provides a detailed sampling of the large spatial variability of the Ebro valley temperature, whereas other available data are unable to capture smaller scale details. In the Ebro valley, the complex topography enhances the importance of small-scale processes in model simulations during summertime. In general, summers are characterized by cyclonic disturbances and low pressure systems. The formation of these cyclones is mostly determined by the thermal heating in the valley. However, these processes can sometimes be modified by the land-sea temperature contrast. Thus, these convective processes require
a spatially dense climate data set to assess the capability of RCMs to represent them. In their study over the Mediterranean region, Gao et al. (2006) found that extreme weather events show a clear topographically-induced fine scale structure over large regions of the Iberian Peninsula. For this reason, there is added value in assessing the magnitude and spatial variations of future climate change processes on fine scales in the Ebro basin.

Figure 1 depicts the spatial distribution of temperature observations. The stations are relatively evenly distributed across the domain, providing not only an adequate sampling of its large spatial heterogeneity of geography and climate, but also a more reliable comparison of model historical data with observations. According to Lenderink (2010), the use of a dense network of observational stations is likely to minimize the uncertainty associated with the model validation outputs. Similarly, Rauscher et al. (2010) and Frei et al. (2003) highlighted the need for a dense observational network for validating climate model outputs.

2.3. Regional Climate Model (RCM) data

A comprehensive set of RCM simulations for Europe is available through the ENSEMBLES project (Van der Linden and Mitchell, 2009). In this work, we used daily $T_{\text{max}}$ data derived from the ENSEMBLES simulations for summers from 1971 to 2100 to assess changes in future climatic conditions. More specifically, daily data for the summer season (JJA) for three different 30-year runs, including a historical period (1971-2000) and two future periods (2021–2050 and 2071-2100), were available for analysis. The daily dataset in ENSEMBLES has an average spatial resolution of $0.25^\circ \times 0.25^\circ$ (approximately 25 km). This computational grid is able to capture sub-regional climate features that may affect temperature variations in the domain (e.g., topography, lee-side effects, and land use/cover changes). Future changes were assessed under the IPCC special report emissions scenarios (SRES) A1B emission scenario, which assumes a mid-range level of greenhouse gas emissions. Table I summarizes the nine models explored in this study and their different experiments. It should be noted that only those ENSEMBLES simulations whose data include the end of the 21st century were used. Further information on the ENSEMBLES and their simulations can also be found at http://ensemblesrt3.dmi.dk/ (Van der Linden and Mitchell, 2009).
As indicated in Table I, the RCM experiments were forced by different GCMs simulation. For example, while CNRM and DMI-ARP were driven by APREGE, DMI-EC and ICTP were forced by ECHAM-3r. Each simulation used different schemes to assimilate large-scale climate, which may raise some significant differences among these simulations (Van der Linden and Mitchell, 2009). Overall, the daily dataset in ENSEMBLES has an average spatial resolution of 0.25° × 0.25° (approximately 25 km). This computational grid is reasonable to capture sub-regional climate features that may affect temperature variations in our domain (e.g., topography, lee-side effects, land use/cover changes). In this study, future changes were assessed under the IPCC SRES A1B emission scenario, which assumes a mid-range level of greenhouse gas emissions.

2.4. RCMs evaluation

The performance of the ENSEMBLES RCMs was evaluated by comparing the model simulations against observational data. The time period for verification of these RCMs was chosen as 1971–2000. To provide an intercomparison of observed versus modeled data, an Inverse Distance Weighting (IDW) interpolation procedure was used to create a grided observational based data set. As the models showed different spatial intervals (refer to Table I), the gridded observational data were delivered on different spatial resolutions so that they exactly overlapped the modeled grid boxes for each particular simulation. The rationale behind this procedure was to fit as closely as possible the rotated latitude-longitude of the modeled grid boxes and maintain consistency between simulated data sets and their corresponding observational grid points.

The performance of each ENSEMBLES model was assessed by using a variety of validation statistics. Combining the skill of a range of statistical estimators provides a more rigorous assessment of the model uncertainty than a single metric used in isolation. Therefore, it was important to assess the RCMs performance by testing their ability to reproduce different aspects of the observed data (e.g., mean, skewness and asymmetry). In this work, we used four accuracy estimators, including bias, the Yule–Kendall (YK) skewness measure (Ferro et al., 2005), the D Willmott’s refined index of agreement (Willmott et al., 2012) and the normalized root mean square error (nRMSE). The validation estimators were calculated independently for each model.
The bias was calculated as a measure of the differences in values between the observed (O) and model (P) data, and given as:

\[ \text{Bias} = N^{-1} \sum_{i=1}^{N} (P_i - O_i) \] ... (1)

where \( N \) is the sample size, \( O \) is the observed value, \( P \) is the model value, and \( i \) is the counter for grid boxes (i.e., the observed and model values for a particular grid box at a definite time step).

The YK estimator was employed as a measure of differences in asymmetry between the observed and model data, suggesting a “perfect” symmetry of data statistical distribution when values are close to 0. The YK was computed following:

\[ YK = \left[ \frac{(pc_{95} - pc_{50}) - (pc_{50} - pc_{5})}{pc_{95} - pc_{5}} \right]_{\text{model}} - \left[ \frac{(pc_{95} - pc_{50}) - (pc_{50} - pc_{5})}{pc_{95} - pc_{5}} \right]_{\text{observed}} \] ... (2)

where \( pc_i \) represents the \( i \)th percentile of temperature distribution.

The refined D Willmott statistic was also computed as a measure of agreement/disagreement between the observed and modeled data, reflecting not only differences in their means, but also differences in their standard deviations. This coefficient is dimensionless as it is bounded between -1 (no skill) and 1 (perfect skill). According to Willmott et al. (2012), the D statistic was given as:

\[ D = 1 - \frac{\sum_{i=1}^{n} |P_i - O_i|}{2 \sum_{i=1}^{n} |O_i - \bar{O}|}, \text{ when } \sum_{i=1}^{n} |P_i - O_i| \leq 2 \sum_{i=1}^{n} |O_i - \bar{O}| \] ... (3)

\[ D = -1, \text{ when } \sum_{i=1}^{n} |P_i - O_i| > 2 \sum_{i=1}^{n} |O_i - \bar{O}| \]
nRMSE is a non-dimensional form of the RMSE, in which RMSE was normalized to the range of the observed data (X). Here, nRMSE was given in percentage (%) and written, as:

\[
nRMSE = \left( \frac{RMSE}{X_{\text{max}} - X_{\text{min}}} \right) \times 100
\]

where \( RMSE = \sqrt{\frac{\sum_{i=1}^{n} (X_{O_i} - X_{\bar{O}})^2}{n}} \)

... (5)

\[
\sum_{i=1}^{n} P_i - O_i > 2 \sum_{i=1}^{n} O_i - \bar{O}
\]

... (4)

2.5. Future changes in regional temperature

Changes in greenhouse gas concentrations will not only impact temperature means, but they will also influence the warm/cold tails of the distribution. While information on changes in the mean temperature is important to understand the variability of temperature, changes in extreme events are more relevant when it comes to economic and societal impacts. Therefore, it is important to explore whether changes in summer T\text{max} are primarily due to the shift of the whole distribution to warmer values or to changes only in the mean temperature. For this reason, changes in the first (mean) and second (standard deviation) moments of summer T\text{max} were assessed. Future changes in the mean and the standard deviation were calculated for each simulation as departures from the long-term mean (1971-2000). In addition to changes in the mean climate, future changes in the corresponding time varying percentiles were also assessed. The aim was to identify whether observed and future changes in the upper and the lower percentiles of temperature distribution were detectable and consistent with those of the temperature mean conditions.

Here, we should emphasize that this work aims to provide a detailed comparison of the individual model projections, which allowed not only for potential uncertainties to be addressed through model inter-comparison, but also for focusing attention on the interrelation and comparison of RCMs driven by different GCMs. We also evaluated climate model output by using an ensemble mean of the nine members examined in this work. The use of the central estimate of models that may differ in their simulated
temperature change and variability is useful as it provides more accurate quantification of various sources of uncertainty in model performance.

In order to examine the statistical distributions and the level of consistency between and within RCM simulations, evaluations were based on the cumulative distribution function. Here, the \( i \)th time varying percentile of a data sample for each 30-year run was simply computed as the value below which \( i\% \) of the data fall. For instance, given 100 occurrences, the 99th percentile (pc99) can be interpreted as an event that likely occurs no more than once per summer. In this research, the \( i \)th percentile magnitude was calculated independently for each ensemble member and for each grid box, where \( i \) relates to the 1, 5, 10, 90, 95, and 99 percentile. This number of percentiles provides insight into changes in different parts of the probability distribution of summer \( T_{\text{max}} \). Lower percentiles (pc1, pc5 and pc10) were used as indicators of changes in the cold tail of temperature distribution, whereas upper percentiles (pc90, pc95 and pc99) give an indication of changes in the warm tail. For summer \( T_{\text{max}} \), changes in the mean, standard deviation and time varying percentiles were assessed for both near recent (1971-2000) and future climates (2021-2050 and 2071-2100) simulations.

An additional aspect that was explored is the extent to which variations in the time varying percentiles are linked to changes in the mean values or in the tails of \( T_{\text{max}} \) distribution. In this regard, it is expected that the inter-annual variation of the time varying percentiles can be linked to changes in the mean when the time series are uncentered, since inter-annual variations of time series decrease when the time series are decentered. Following the procedure detailed by Beniston and Stephenson (2004), the mean and time varying percentile time series were uncentered for both the current (1971-2000) and future (2021-2050 and 2071-2100) climates by subtracting the median (i.e., pc50) from each time series. The magnitudes of the median and the mean are expected to be very close because temperature data typically have a symmetric (Gaussian) distribution (Wilks, 1995). The Pearson correlation coefficient \( (r) \) was then computed between the uncentered time varying percentile time series and the uncentered mean time series for each ensemble member and for each grid box. The statistical significance of this dependency was assessed at the 95% confidence level \( (p<0.05) \) using Kendall’s tau test. This statistic makes no assumptions about the distribution of the time series residuals and is robust to the effect of outliers in the
series. As expected, higher correlation suggests a consensus among long-term changes in the mean and the corresponding percentiles in terms of the magnitude and/or the direction (sign). Conversely, lower correlation provides an indication of disagreement.

3. Results

3.1. Model validation results

In this study, a set of statistical measures is used to analyze the ability of an ensemble of multiple RCMs to simulate observed climate over the Ebro valley, with an aim to assess how well individual models reproduce the observed climate between 1971 and 2000. Box plots summarizing the validation results are given in Figure 2. As illustrated in Figure 2-a, the majority of the models underestimate summer (JJA) maximum temperature across much of the study domain. The CNRM, MPI, KNMI and DMI-ARP members show the lowest values, with a mean bias of 0.9, -1.1, -1.7 and 1.9°C, respectively. In contrast, other members exhibit poor performance, with DMI-EC (-4.6°C), DMI-BCM (-6.8º C), ICTP (-3ºC), METO (4.1ºC) and SMHIRCA (-4ºC). In terms of the YK coefficient (panel b), the CNRM, MPI, KNMI and DMI-ARP models also gave the best validation results, with near-to-zero values (CNRM=-0.01ºC, MPI=0.01ºC and KNMI=0º C). Similarly, Figure 2-c suggests higher D Willmott statistic for the CNRM, MPI, KNMI and DMI-ARP simulations, with values of 0.34, 0.33, 0.29 and 0.45, respectively. The nRMSE results also reflect agreement with other accuracy estimators, with the best performance found for DMI-ARP (19.3%), MPI (20.2%), CNRM (21.1%) and KNMI (22.3%). Other models give values further away from the ideal zero value.

While the previous accuracy statistics focused mainly on the statistical properties of the time series (e.g., mean, variance and skewness), it is also important to verify that the models are skillful in reproducing the spatial patterns of the observed temperature. Therefore, the performance of the models in terms of capturing the regional features of temperature was also assessed by comparing the simulated summer climatology with gridded observational data for the control period (1971–2000), with results shown in Figure 3 for all models and for an ensemble average. Interestingly, the simulations which show better performance in terms of the statistical measures (e.g., CNRM, DMI-ARP and MPI) are also more skillful in reproducing the spatial patterns of the observed temperature (left
panels). These simulations show more or less consistent spatial patterns of temperature, with a distinct north-south gradient across the Ebro valley. The largest changes occur over the central parts of the Ebro catchment, whereas lower temperatures are recorded over northern areas and the mountainous belts (e.g., the Pyrenees to the north and the Iberian system to the southwest).

Inspection of Figure 3 (left panels) shows some interesting results. First, the models are adequately skillful in reproducing regional temperature patterns across the domain (roughly 85,000 km²), even though a significant part of the temperature patterns in this region are driven by local topographical influences. The finding is interesting because the noise associated with temperatures is expected to be greater at local and regional scales than at larger continental or global scales. Second, the good agreement between the simulated and observed changes corresponds to RCM experiments, which are forced by the same GCM experiment. In particular, both CNRM and DMI-ARP are driven by the ARPEGE experiment. Likewise, the models which are forced by the ECHAM5-r3 experiment (i.e., DMI-EC, DMI-BCM and MPI) show identical spatial patterns of maximum temperature during the summer. Figure 3 also reveals that, in contrast to the mean summer $T_{\text{max}}$ temperature, the RCM simulations have little skill in reproducing the spatial patterns of standard deviation. As illustrated, all models fail to reproduce the spatial patterns and the magnitude of standard deviation. Apparently, standard deviation is greatly overestimated across the whole domain. For the coefficient of variation (CV), the models give both the correct sign and the spatial variability over the study domain. In particular, the models simulate higher values in the mountainous areas, with lower values recorded over central parts of the study area.

### 3.2. Future changes in temperature statistics

After evaluating the model performance for the control run (1971-2000), future changes in summer $T_{\text{max}}$ were evaluated using each individual member as well as the ensemble mean. The potential impacts of recent climate change in the study domain were assessed by comparing the differences between the control run and future simulated climate. Figure 4 shows summer $T_{\text{max}}$ anomalies for the period 2021-2050. As can be seen, summer $T_{\text{max}}$ exhibits a positive anomaly in the different models and regions. Spatially, the patterns associated with the best-validated members indicate that this anomaly is expected to be
much stronger over northeastern parts including the Pyrenees, the Iberian system southwest and to a lesser extent in the central Ebro valley and areas close to the Mediterranean and the Cantabian coast. In the central Ebro valley, summer T\textsubscript{max} may increase by 2-3\textdegree C from 1971-2000 to 2021-2050. A comparison between the members indicates that the increase in summer T\textsubscript{max} is comparable among most of the simulations, suggesting an ensemble mean range of 0.7\textdegree C to 3.4\textdegree C. Figure 4 (middle panels) shows future changes in the standard deviation of summer T\textsubscript{max}. The standard deviation of summer T\textsubscript{max} does not show clear spatial differences from 2021 to 2050. However, the standard deviation of some simulations (e.g., DMI-BCM and SMHIRCA) increases in areas of low altitude and decreases in areas of high latitude. For the coefficient of variation, all models reveal a distinct SE-NW gradient, with increasing variability occurring close to the Cantabrian Sea.

Figure 5 illustrates summer T\textsubscript{max} changes during the period from 2071 to 2100, computed as departures from the control period. Summer T\textsubscript{max} will get warmer in the far future with rates markedly higher than in the near future (CNRM= 4.4\textdegree C, DMI-ARP= 3.9\textdegree C). The ensemble average suggests a projected increase of 4.6\textdegree C, with maxima in the Pyrenees (approximately 6\textdegree C) and minima in areas close to the coast (approximately 3.2\textdegree C). However, some simulations show increases, which are markedly higher than the ensemble mean (METO= 6.1\textdegree C). This reveals that the spread among the models is markedly larger from 2071 to 2100, compared with the 2021-2050 period. The range is also high for the standard deviation, where some simulations suggest an increase relative to the control period (e.g., CNRM= 0.3\textdegree C, ICTP= 0.4\textdegree C, KNMI= 0.2\textdegree C, MPI=0.7\textdegree C and SMHIRCA= 0.4\textdegree C), while other models tend to show a decreasing standard deviation (e.g., DMI-ARP= -1.3\textdegree C, DMI-EC= -2.9\textdegree C and METO= -0.9\textdegree C). Averaged across all individual members, the standard deviation will decrease by -0.2 to -1.1\textdegree C from 2071 to 2100. This suggests that summer T\textsubscript{max} tends to deviate less from the mean in the future, which may imply that the shift toward warmer conditions in the future will not be accompanied with an increase in the inter-annual variability of temperature.

Given the projected increase in the mean temperature and the corresponding decrease in standard deviation, the results suggest a shift in the distribution of summertime T\textsubscript{min} under the A1B emission scenario, but with a tendency to have more intense and frequent warm summer extremes over the region at the end of the century. Figure 6 indicates that for the
majority of the simulations, summer Tmax tends to shift with positively skewed and more widening and flattening temperature distributions, relative to the control period. As illustrated, the warm tail of the majority of the models is likely to be positively skewed, suggesting a decrease in the cold events in the future. The only exception corresponds to those models, which showed a decrease in the standard deviation (e.g., DMI-ARP, DMI-EC and METO). Such results provide an indication of a rapid shift in the warm tail of summer Tmax distribution, suggesting more warm extremes. For the coefficient of variation, the projected changes are relatively homogeneous over the entire domain. All solutions and the ensemble average produce a similar pattern with a clear regional gradient along a SE-NW axis, which is consistent with the spatial structure found in the 2021-2050 period.

3.3. Future changes in the time varying percentiles

Results comparing changes in the magnitude of the percentiles of summer Tmax relative to the control period, are summarized in Table II. All percentiles show a steady increase in their magnitude, indicating that the current warming is projected to continue during the 21st century. Figure 7 illustrates changes in the percentiles over the period 2021-2050 relative to changes in the mean of the total distribution for the same period. The results show that the rates of change across the percentiles are relatively stable for some of the members, particularly for DMI-EC, KNMI and MPI. In some cases, the lower percentiles exhibit greater rates of change compared with upper percentiles (see CNRM and METO), while in others the reverse is true: upper percentiles show more changes than lower percentiles (e.g., DMI-ARP). Taken together, the results illustrate that there is a high degree of variability between the models in terms of the association between changes in the mean and those of the percentiles.

Figure 8 shows changes in the upper and lower percentiles for the period 2071-2100. As depicted, the rates of changes for the upper percentiles are higher than those of the lower percentiles. This is clearly evident for all models, except for METO. One representative example is the pc99, which moves towards more extremes, with an average rate increase of 34.7% for DMI-EC, 26.3% for DMI-BCM, 36.1% for MPI and 27% for SMHIRCA. On the other hand, lower percentiles show smaller rate changes. For instance, the rates of changes for the pc5 are 18.8% for DMI-EC, 4.9% for DMI-BCM, 23.5% for MPI and 14.9% for SMHIRCA. This finding suggests that future changes in the upper and lower percentiles
during summertime will not necessarily be consistent with the rates of change in the mean temperature. Figure 8 also informs that the percentiles will exhibit a systematic change in their magnitude as a function of the percentile order. More specifically, changes in the magnitude of the percentiles generally increase towards the upper end of the probability distribution. Changes in the 99th percentiles are larger than changes in other percentiles. Thus, it is a robust signal across the entire ensemble that very hot extremes will increase at rates higher than those of moderate extremes. Figure 8 also reveals that the spread among the models in far future simulations is larger than that of the near future projections. As shown in Table II, the projected increase in pc99 varies from 0.5°C (DMI-BCM) to 2.2°C (METO) in the period 2021-2050, compared with a range from 3.1°C (DMI-BCM) to 7.4°C (MPI) at the latter half of the 21st century.

In order to assess whether future changes in the mean temperature are consistent with changes in the corresponding time-varying percentiles, we computed Pearson correlation coefficients between inter-annual variations in the mean of summer Tmax and those of the time varying percentiles. For each individual member, the correlation was first computed for each pixel and then averaged for the whole domain. Figure 9 summarizes the results for the two future time periods (2021-2050 and 2071-2100). As depicted, the projected inter-annual variations in average summer Tmax are more consistent with variations in the percentiles. The correlation coefficients are generally above 0.5, which is statistically significant at the 95% level. Nonetheless, the correlation coefficients are found much stronger for moderate extremes (i.e., pc10 and pc90) than for extreme percentiles (i.e., pc1, pc5, pc95 and pc99): a finding that is evident in all models. For example, changes in pc10 and pc90 time series correlate well with changes in the mean, with coefficients generally larger than 0.8. In contrast, the coefficient values are weaker for the most extreme percentiles (i.e., pc1 and pc99). This finding implies that future changes in the mean can be better linked with variations in the upper percentiles (particularly pc90 and pc95) than with lower percentiles.

4. Discussion

4.1. Validation outputs

A set of statistical measures (e.g., bias, YK, Willmott's D refined index and nRMSE) were used to validate the outputs from nine RCMs from the ENSEMBLES project over the Ebro
valley from 1971 to 2000, providing insights into the differences in the mean, variance and
asymmetry of temperature between the observed and simulated climate. Model validation
results indicate that there are significant differences among the models in terms of their
ability to reproduce the near-present climate. An inter-model comparison suggests that a
suite of RCMs analyzed in this study (i.e., CNRM and DMI-ARP) were able to reasonably
reproduce the observed distribution and variability of summer $T_{\text{max}}$. However, the results
also suggest that other RCMs may not adequately simulate future climate change signals, as
they show some error in present day reproductions of climate. Interestingly, results indicate
that the models which better simulate summer $T_{\text{max}}$ are driven by the same GCM. From a
modeling point of view, this can be expected given that these models are climatically and
physically consistent, with a more or less similar ability to capture fluctuations in energy
budgets during the daytime. The models with the worst performance (e.g., KNMI, ICTP
and SMHIRCA) were those that used a wider boundary relaxation zone for air advection
(Van der Linden and Mitchell, 2009). There are features of natural variability that are likely
showing up in the GCM simulations, which may be passed through the boundaries to
RCMs, implying that these simulations maintain the large-scale circulation (e.g., SST)
inside the RCM domain closer to the forcing fields. This may suggest that the biases found
for some models are driven by the GCMs and not induced by the RCMs themselves.

In summer, some large-scale features that differ significantly from the observations (e.g.,
long-term persistence of some atmospheric circulation modes) might be represented during
the control period of these models, causing poor performance. This feature should be seen
in the context that the study domain is quite small (approximately 85,000km$^2$), and
therefore temperature is not expected to vary greatly as a consequence of large-scale
configurations during summer. While local thermal heating is a key driver of summer
temperature, the influence of large-scale circulation can be modulated by topography and/or
land-sea thermal contrasts. Unfortunately, there is a lack of available mid-troposphere
fields in the ENSEMBLES archive to assess the contribution of the large-scale circulation
to the detectable bias. Over Iberia, Jerez et al. (2013) found that the sign of the projected
climate change varies largely depending on the choice of the model physics. Accordingly, it
has been suggested that an ensemble combining numerous members may work better at
representing summer temperature conditions in the region than one or even a combination
of a few models that perform well. The multi-model ensemble approach has been favored
in many studies rather than relying on individual models (e.g., Christensen et al., 2007; Evans, 2009; Arritt and Rummukainen, 2010).

In Figure 10, we attempted to rank the models according to their respective performances by combining the results from the four validation statistics used in this work. As our validation statistics (bias, YK, D and nRMSE) were given in quantities with different units, we first normalized the output values of each of these statistics, so that they had a mean of 0 and standard deviation of 1. This procedure constrained the largest weight that could possibly be given to some validation statistics (e.g., the bias), which may vary mostly between the models. Then, the overall skill of each model was computed at each grid box as an arithmetic average of the normalized statistics. Next, the ensemble members were objectively ranked for each grid box on the basis of their overall skill, where skill score closer to 0 refers to the model with the best performance, and scores furthest from 0 indicate the model with the worst performance. The results indicate that no unique model was classified as the best over the whole domain. As depicted in Figure 10, both CNRM and DMI-ARP were identified as the best in almost 42.7% of the grid boxes across the whole domain. Averaging their statistics for the whole domain, some models (e.g., KNMI and MPI) were found skillful in reproducing some of the temperature features. In contrast, when the skill of these models was assessed at a more local scale (i.e., grid box), the results vary significantly. This implies that the models can, on average, show good agreement with observations, whereas they probably introduce large local differences. This finding illustrates again that an ensemble mean can guarantee a better prediction of future climate for the entire territory.

Inspection of the RCM simulations reveals that the models tend to underestimate summer Tmax over large parts of the region. The under-prediction of summer Tmax might, to some extent, be explained by inadequate representations of convective parameterizations or components of the land-surface schemes, such as soil moisture schemes and land-atmosphere flux parameterizations (Mueller et al., 2013; Xue et al., 2014). Many authors (e.g., Giorgi and Marinucci, 1996; Rio et al., 2009) attributed much of the bias in climate models during summertime to a lack of adequate treatment of cumulus convection; a process that plays a key role in temperature variations during this season. In southern Europe, Moberg and Jones (2004) attributed bias in mean temperature to poor simulation of
soil moisture. Over the same domain, Hohenegger and Vidale (2005) also attributed the bias of summer temperature to the specification of aerosols. Another possible explanation of temperature underestimation during summer months is the grid resolution of the models used in this work, which might not be sufficiently fine to simulate processes associating with summer convective activities. Modeling summertime temperature may be constrained by very local conditions near the land surface, such as local topography, vegetation cover and wind velocity, all of which are difficult to resolve by the current spatial resolution of RCMs (25×25 km) (Evans and McCabe, 2013).

4.2. Future changes in summer T_max

In this work, future changes in summer T_max were assessed for the 21st century using each individual member as well as an ensemble average. Analysis of the simulated temperature datasets provides strong evidence that the substantial warming found during the latter half of the 20th century (El Kenawy et al., 2012) is set to continue in the future. This warming will be markedly larger at the end of the 21st century. From 2071 to 2100, summer T_max is expected to increase by 3.2 to 6.1ºC, while it will warm by 0.7 to 3.7ºC from 2021 to 2050. In terms of magnitude, the values of the warming estimated here are in the midrange of previously estimated assessments over the Iberian Peninsula and the Mediterranean region (e.g., Meehl et al., 2007; Giorgi and Lionello, 2008; Nogués-Bravo et al., 2008; Argueso, 2011). Meehl et al. (2007) found that the global mean surface air temperature is projected to rise by 1.7- 4.4ºC by the late 21st century. Over the Mediterranean, Giorgi and Lionello (2008) projected an increase in summer T_max from 2081 to 2100 by 4.6ºC. In the Iberian Peninsula. Argueso (2011) indicated that summer T_max is expected to increase under the A1B emission scenario by more than 4ºC from 2070 to 2099. Déqué et al. (2007) also projected a dramatic warming in summer temperature over the Iberian Peninsula from 2071 to 2100, with values ranging between 4 and at least 7ºC. The projected warming can partially be explained by the projected deficit in soil moisture in the future. Recently, many studies linked the occurrence of heatwaves over Europe with the influences of land surface processes, including soil moisture (Miralles et al., 2012, 2014). Other previous work (e.g., Kjellström et al., 2007; Jerez et al., 2012; Stegehuis et al., 2013) also found that the soils may become drier under future warming and available energy is more likely to be used for heating the air (sensible heat).
Spatially, there is a general consensus among the models in the projected summer $T_{\text{max}}$ changes of: (1) a positive anomaly of temperature over high altitudes, particularly in the Pyrenees and the Iberian system; and (2) a slight positive temperature anomaly over the central parts of the Ebro valley and along coastal areas. These regional differences fit well with recent climate model projections over Iberia (e.g., López- Moreno et al., 2008; Argueso et al., 2011; Jerez et al., 2012). For example, Jerez et al. (2012) projected an increase in summer temperature over mainland areas of the Iberian Peninsula. In their study on the Pyrenees, López- Moreno et al. (2008) reported a rapid increase in mean temperature over land grid points relative to the Mediterranean grids. The agreement between the models on the strong warming over mountainous areas is interesting because temperature is largely influenced by topographic complexity (e.g., peaks and slopes) in these areas. On the local scale, the orography is a dominant factor in the development of cloud systems during summer, which leads to strong temperature differences between uplands and lowlands and/or between slopes with different exposures.

4.3. Future changes in summer time varying percentiles

While results on changes in mean temperature are of considerable interest, the impacts of future change on society are more likely associated with extreme climatic events. A comparison between the role of changes in the mean temperature of summer $T_{\text{max}}$ and their corresponding time varying percentiles during the 21st century reveals some interesting results. Results indicate that the simulated changes in both upper and lower percentiles of temperature distributions do not show the same rates of change in the mean. Therefore, we cannot rely solely on changes in the mean to infer changes in the corresponding time varying percentiles. Overall, there is consensus among models on the high likelihood of increases in the upper percentiles of summer $T_{\text{max}}$, especially at the end of this century. This suggests that impacts of more intense and frequent warm extremes could be severe in the region in the future. Numerous modeling studies over the pan-European continent indicate that upper and lower percentiles of temperature distribution tend to increase at different rates, compared with future changes in the mean temperature (e.g., Clark et al 2006; Kjellström et al., 2007; Hertig et al., 2010). Hertig et al. (2010) projected an increase in the magnitude of the upper percentiles of summer maximum temperature over the western Mediterranean in the 21st century. They indicated that changes in the pe90 of summer $T_{\text{max}}$
will likely not follow the same rates of change as the mean values. Kjellström et al. (2007) found that the simulated changes in extreme temperatures during summer over most of the European sub-regions would be larger than changes in the mean.

A more detailed inspection of the association between changes in the mean on the one hand and changes in the time varying percentiles on the other hand reveals that the greatest association was found for more moderate percentiles (i.e., pc10 and pc90) than for very extreme percentiles (i.e., pc1 and pc99). This can be explained by the fact that the less extreme percentiles represent a larger part of the distribution and are therefore more important for the mean temperature than are the most extreme values. Another possible explanation can be related to the influence of projected reductions in soil moisture, which is more significant at moderate percentiles as the climate becomes drier. Conversely, the most extreme percentiles are already accompanied by dry climates. Clark et al. (2006) confirmed this finding in their assessment of projected increases in summer heat extremes in the Northern Hemisphere.

Results on future changes in the percentiles of summer $T_{\text{max}}$ suggest a larger spread between RCMs at the end of the century, which is clearly revealed in the larger spatial variability of summer $T_{\text{max}}$ percentiles (Table II). While all models display increases in both the mean and magnitudes of the time varying percentiles for summer $T_{\text{max}}$ in the future, standard deviations are projected to decrease in future projections (Figures 4 and 5). The majority of the models tend to show a decrease in the standard deviation for summer $T_{\text{max}}$ in far future projections, implying that the daily temperatures are expected to become less variable, in agreement with van Oldenborgh et al. (2009) and Min et al., (2013) for north-western Europe. Other studies reveal that summer temperature variability in the Iberian Peninsula exhibits different behaviour, relative to other parts in Europe (see for example, Della-Marta et al., 2007; Parey et al., 2010). For example, Della-Marta et al. (2007) identified that although the Iberian Peninsula experienced more extreme heat waves during the last decades, its summer temperature variability decreased by $-7\pm3\%$. The results from this current analysis confirm that this feature will continue in the future.

5. Conclusions
An evaluation of the ability of a suite of high resolution RCMs to capture the main statistical features (i.e., the mean, skewness and asymmetry) of the observed regional climate in the Ebro basin of northeast Spain was undertaken. Although the simulated summer $T_{\text{max}}$ reproduced the observed spatial variation reasonably well, a tendency in most of the models to underestimate $T_{\text{max}}$ during summer was noted. An inter-model comparison indicated that not all models were consistent in describing the present climate in terms of the mean, skewness and asymmetry. It was determined that a more balanced assessment of future climate for the Ebro basin can be achieved by using an ensemble mean rather than relying on individual model reproductions.

To account for possible changes in the distribution of summer $T_{\text{max}}$, we assessed future changes in climate in terms of changes in the mean and standard deviation as well as through a set of time-varying time varying percentiles (pc1, pc5, pc10, pc90, pc95 and pc99). Analysis of the simulated temperature, driven by the A1B emission scenario, suggested that the regional warming in both the mean and extreme events is expected to continue during the 21st century, with a stronger increase in the latter half of the 21st century. The magnitude of climate change signal was found to be greater for upper percentiles of summer $T_{\text{max}}$ than for the mean and lower percentiles, suggesting a tendency towards more intense and frequent warm extremes in the future. Spatially, all models agree on a stronger positive anomaly of summer $T_{\text{max}}$ over high mountains compared with mainland and coastal area.

Understanding the impacts of future climate change on the Ebro valley, a region characterized by complex climatological and topographical features, is important for environmental, hydrological, agricultural and socioeconomic applications. The analysis presented here provides evidence that the current substantial warming over the studied domain will continue during the 21st century. The projected warm summers and the spatial structure of future changes highlight the need to adapt future policy and development plans in the region to address these local and regional scale responses.

**Acknowledgments**

We would like to thank the Spanish Meteorological State Agency (AEMET) for providing the database used in this study. Research reported in this publication has been supported by
the King Abdullah University of Science and Technology (KAUST), Saudi Arabia. This work has also been supported by the research project CGL2011-27574-CO2-02 financed by the Spanish Commission of Science and Technology and FEDER and “Demonstration and validation of innovative methodology for regional climate change adaptation in the Mediterranean area (LIFE MEDACC)” financed by the LIFE programme of the European Commission. Finally, the ENSEMBLES data used in this work were funded by the EU FP6 Integrated Project ENSEMBLES whose support is gratefully acknowledged.

References


Domínguez, M., R. Romera, E. Sánchez, L. Fita and others (2013). Present-climate precipitation and temperature extremes over Spain from a set of high resolution RCMs. Climate Research 58:149-164.


Evans, J. P. and M. F. McCabe (2013). Effect of model resolution on a regional climate model simulation over southeast Australia Climate Research, 56, 131-145.


methods for Europe and European regions—assessing their performance with respect to extreme temperature and precipitation events. CRU RP11, University of East Anglia.


increasing drought severity caused by temperature rise in southern Europe.


of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs.
Climatic Change 62(1-3): 189-216.

dynamical downscaling in intraseasonal to seasonal simulation/prediction and major
Table I. Summary of RCMs simulations used in the study with their corresponding resolutions.

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Institute</th>
<th>Driving data</th>
<th>RCM</th>
<th>Resolution (Lat. by Lon.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNRM</td>
<td>Météo France (France)</td>
<td>ARPEGE</td>
<td>Arpège</td>
<td>0.15° by 0.35°</td>
</tr>
<tr>
<td>DMI-ARP</td>
<td>Danish Meteorological Institute (Denmark)</td>
<td>ARPEGE</td>
<td>HIRHAM5</td>
<td>0.17° by 0.32°</td>
</tr>
<tr>
<td>DMI-EC</td>
<td>Danish Meteorological Institute (Denmark)</td>
<td>ECHAM5-r3</td>
<td>HIRHAM5</td>
<td>0.17° by 0.32°</td>
</tr>
<tr>
<td>DMI-BCM</td>
<td>Danish Meteorological Institute (Denmark)</td>
<td>BCM</td>
<td>HIRHAM5</td>
<td>0.17° by 0.32°</td>
</tr>
<tr>
<td>ICTP</td>
<td>The Abdus Salam Intl. Center for Theoretical Physics (Italy)</td>
<td>ECHAM5-r3</td>
<td>RegCM3</td>
<td>0.18° by 0.30°</td>
</tr>
<tr>
<td>KNMI</td>
<td>Koninklijk Nederlands Meteorologisch Instituut, (Netherlands)</td>
<td>ECHAM5</td>
<td>RACMO2</td>
<td>0.17° by 0.32°</td>
</tr>
<tr>
<td>METO</td>
<td>Norwegian Meteorological Institute (Norway)</td>
<td>HadCM3Q16</td>
<td>HIRHAM</td>
<td>0.17° by 0.32°</td>
</tr>
<tr>
<td>MPI</td>
<td>Max-Planck-Institut für Meteorologie (Germany)</td>
<td>ECHAM5-r3</td>
<td>REMO</td>
<td>0.17° by 0.32°</td>
</tr>
<tr>
<td>SMHIRCA</td>
<td>Swedish Meteorological and Hydrological Institute (Sweden)</td>
<td>BCM</td>
<td>RCA3</td>
<td>0.17° by 0.32°</td>
</tr>
</tbody>
</table>
Table II: Future changes in the percentiles of summer $T_{\text{max}}$ during the near (2021-2050: I) and far future (2071-2100: II). Changes are calculated relative to the control period (1971-2000) and given in °C.

<table>
<thead>
<tr>
<th></th>
<th>pc1</th>
<th>pc5</th>
<th>pc10</th>
<th>pc90</th>
<th>pc95</th>
<th>pc99</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>II</td>
<td>I</td>
<td>II</td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>CNRM</td>
<td>1.8</td>
<td>3.1</td>
<td>2.4</td>
<td>4.3</td>
<td>2.3</td>
<td>4.5</td>
</tr>
<tr>
<td>DMI-ARP</td>
<td>1.4</td>
<td>3.2</td>
<td>1.3</td>
<td>3.0</td>
<td>1.6</td>
<td>3.3</td>
</tr>
<tr>
<td>DMI-EC</td>
<td>1.5</td>
<td>3.2</td>
<td>1.4</td>
<td>2.9</td>
<td>1.4</td>
<td>3.1</td>
</tr>
<tr>
<td>DMI-BCM</td>
<td>0.6</td>
<td>2.3</td>
<td>0.9</td>
<td>2.2</td>
<td>0.6</td>
<td>2.0</td>
</tr>
<tr>
<td>ICTP</td>
<td>0.5</td>
<td>3.7</td>
<td>0.5</td>
<td>3.6</td>
<td>0.8</td>
<td>3.9</td>
</tr>
<tr>
<td>KNMI</td>
<td>2.0</td>
<td>3.8</td>
<td>1.8</td>
<td>4.7</td>
<td>1.3</td>
<td>4.8</td>
</tr>
<tr>
<td>METNO</td>
<td>4.8</td>
<td>7.9</td>
<td>3.2</td>
<td>7.4</td>
<td>2.3</td>
<td>6.5</td>
</tr>
<tr>
<td>MPI</td>
<td>1.5</td>
<td>4.2</td>
<td>1.4</td>
<td>4.1</td>
<td>1.4</td>
<td>4.6</td>
</tr>
<tr>
<td>SMHIRCA</td>
<td>0.3</td>
<td>2.6</td>
<td>0.5</td>
<td>3.0</td>
<td>0.4</td>
<td>3.0</td>
</tr>
<tr>
<td>Ensemble Mean</td>
<td>1.6</td>
<td>3.8</td>
<td>1.5</td>
<td>3.9</td>
<td>1.3</td>
<td>4.0</td>
</tr>
</tbody>
</table>
Figure 1: Topography of the Ebro Valley and location of the meteorological stations with continuous daily records of surface air temperature for the period 1971-2000.
Figure 2: Box plots summarizing validation results computed for all grid boxes in the study domain from 1971 to 2000. The blue central line represents the mean, the black central line indicates the median, top and bottom of boxes show the 75th and 25th percentile, respectively, and upper and lower whiskers represent the 90th and 10th percentile, respectively.
Figure 3: Spatial distribution of the mean (left), standard deviation (middle) and coefficient of variance (right) of summer $T_{\text{max}}$, calculated for both the observed and model data for the control period (1971-2000). The ensemble mean is calculated as the arithmetic average of the nine RCMs used in this study.
Figure 4: Spatial distribution of summer $T_{\text{max}}$ anomaly during the period 2021-2050. The anomalies are calculated for the mean (left), standard deviation (middle), and coefficient of variation (right), and defined as departures from the long-term mean (1971-2000) of the model data (i.e., 2021-2050 minus control period). The legend representing values range was not homogenized for all panels to improve the final look of the graph given that some models showed high range.
Figure 5: Spatial distribution of summer $T_{\text{max}}$ anomaly during the period 2071-2100. The anomalies are calculated for the mean (left), standard deviation (middle), and coefficient of variation (right), and defined as departures from the long-term mean (1971-2000) of the model data (i.e., 2071-2100 minus control period). The legend representing values range was not homogenized for all panels to improve the final look of the graph given that some models showed high range.
Figure 6: Cumulative distribution of summertime maximum temperature for the control period (1971-2000) against the near (2021-2050) and far (2071-2100) future simulations.
Figure 7: Box plots representing changes in the values of the percentiles for summer $T_{\text{max}}$ from 2021 to 2050, relative to changes in the mean of the whole distribution. The rates of changes were calculated, following $\left(\frac{PC_{\text{future}} - PC_{\text{control}}}{Mean_{\text{future}} - Mean_{\text{control}}}\right) \times 100$. The dotted line refers to identical changes (i.e., no differences) between changes in the mean and changes in the percentiles.
Figure 8: Box plots representing changes in the values of the percentiles for summer $T_{\text{max}}$ from 2071 to 2100, relative to changes in the mean of the whole distribution. The rates of changes were calculated, following $\frac{(PC_{\text{future}} - PC_{\text{control}}) - (Mean_{\text{future}} - Mean_{\text{control}})}{Mean_{\text{future}} - Mean_{\text{control}}} \times 100$. The dotted line refers to identical changes (i.e., no differences) between changes in the mean and changes in the percentiles.
**Figure 9**: Area average Pearson correlation coefficient (r) between time series of the time varying percentiles and those of summer T_max. The dependency is calculated for the 2021-2050 (left half of each panel) and 2071-2100 (right half of each panel). The calculation was first computed for each pixel in the study domain and then averaged for the whole domain. Pearson coefficients above 0.36 are statistically significant at the 95% level.
Figure 10. Frequency distribution of the percentage of grid boxes assigned to each model, according to the overall skill.