Climatic Change (2007) 81:249–265
DOI 10.1007/s10584-006-9220-5

Modelling daily temperature extremes: recent climate and future changes over Europe

Erik Kjellström · Lars Bärring · Daniela Jacob · Richard Jones · Geert Lenderink · Christoph Schär

Received: 15 February 2005 / Accepted: 17 October 2006 / Published online: 17 March 2007

Abstract Probability distributions of daily maximum and minimum temperatures in a suite of ten RCMs are investigated for (1) biases compared to observations in the present day climate and (2) climate change signals compared to the simulated present day climate. The simulated inter-model differences and climate changes are also compared to the observed natural variability as reflected in some very long instrumental records. All models have been forced with driving conditions from the same global model and run for both a control period and a future scenario period following the A2 emission scenario from IPCC. We find that the bias in the fifth percentile of daily minimum temperatures in winter and at the 95th percentile of daily maximum temperature during summer is smaller than 3 (±5°C) when averaged over most (all) European sub-regions. The simulated changes in extreme temperatures both in summer and winter are larger than changes in the median for large areas. Differences between models are larger for the extremes than for mean temperatures. A comparison with historical data shows that the spread in model predicted changes in extreme temperatures is larger than the natural variability during the last centuries.

E. Kjellström (✉) · L. Bärring
Rossby Centre, SMHI, SE-60176 Norrköping, Sweden
e-mail: Erik.Kjellstrom@smhi.se

D. Jacob
Max Planck Institute for Meteorology, Hamburg, Germany

R. Jones
Met Office Hadley Centre (Reading Unit), Reading, UK

G. Lenderink
KNMI, de Bilt, The Netherlands

C. Schär
Atmospheric and Climate Science, ETH, Zurich, Switzerland

L. Bärring
GeoBiosphere Science Centre, Lund University, Lund, Sweden

1 Introduction

Regional climate models (RCMs) have the potential to provide detailed information not only on mean conditions but also on extremes (Beniston et al. 2007). Given quasi-observed lateral boundary conditions, i.e. from reanalysis experiments, RCMs have been shown to realistically simulate variability of many climate parameters on different temporal scales (Giorgi et al. 2001). When the RCMs are forced with boundary conditions from GCMs, however, these boundary conditions often introduce systematic biases in the simulation of the present climate (Noguer et al. 1998). Some of the systematic biases are amplified when looking into more extreme events like maximum and minimum temperatures (e.g. Moberg and Jones 2004) though this is not always the case, e.g. for percentage errors in precipitation extremes (Buonomo et al. 2007). In the present study, variability and scenario changes in daily maximum and minimum temperatures in a suite of ten RCMs taking part in the European project PRUDENCE (Christensen et al. 2007) are investigated. The RCMs were driven by boundary conditions from the same GCM scenario.

Changes in extreme temperatures over long time series have been described by Yan et al. (2002) using data from ten stations in Europe and China. They conclude that cold extremes have been decreasing and warm extremes increasing during recent decades, but also that there have been earlier changes in these extremes. In an analysis of more than 100 European station records for the second half of the twentieth century, Klein Tank and Können (2003) arrived at a similar conclusion, namely that the cold extremes have been decreasing and the warm extremes increasing during the last quarter of the twentieth century. Here we relate the simulated changes not just to model and inter-model variability but also to the natural variability as reflected in some very long instrumental records.

In a previous study, Kjellström (2004) investigated daily variability of daily mean 2m-temperature and its changes under changed climatic conditions. Here, we look into daily maximum and minimum 2m-temperatures in climate simulations over Europe, including their upper and lower percentiles relating to the occurrence of heatwaves and cold spells, respectively. Changes of these parameters could also be affected by changes in interannual climate variability. It has been suggested that in a future climate the summer interannual variability of surface temperatures might increase as an effect of changes related to land-surface processes (Schär et al. 2004). Most of the models used in the PRUDENCE consortium show indeed some increase in interannual temperature variability during the summer season, but there are considerable inter-model differences regarding the amplitude, geographical location and seasonal timing of the effect (Lenderink et al. 2007; Vidale et al. 2007).

2 Methods and data

2.1 Models

We use results from ten RCMs, nine from PRUDENCE and an additional one run by MetNo (HIRHAM-NO, Hanssen-Bauer et al. 2003), see Jacob et al. (2007) and Déqué et al. (2007) for details. The lateral boundary conditions for all RCMs were provided by the GCM HadAM3H (see Buonomo et al. 2007 for details) and the lower (sea-surface) boundary condition for this were taken from observations and HadCM3 (see Rowell 2005 for details). All models have been run for a control period (CTRL), 1961–1990, and a future scenario period (A2), 2071–2100, following the A2 emission scenario from IPCC (Nakićenović et al. 2000). We study daily maximum ($T_{2\text{max}}$) and minimum ($T_{2\text{min}}$) surface air temperatures as well as the daily mean temperature ($T_{2\text{m}}$). The $T_{2\text{max}}$ and $T_{2\text{min}}$ variables are defined as the maximum and minimum, respectively, of all time-step calculations for a day, which is defined as the time between 00 and 00 UTC. The daily mean is simply the average over these time-steps. The length of the time-steps is generally about 5 min in the models except in RACMO (12 min) and RCAO (36 min).

2.2 Observational station data

We compare model data to observational data from the European Climate Assessment (ECA; Klein Tank et al. 2002). This dataset consists of data from about 200 stations covering much of Europe. Here, we look at the time period 1961–1990. We require that station series should have data for at least 90% of the days in that time period, leaving 147 stations with a geographical distribution as shown in Fig. 1.

In addition to the ECA dataset we also use long-term data from seven stations from the IMPROVE dataset (Camuffo and Jones 2002) that are also shown in Fig. 1. These stations report temperature for the last few centuries reaching back into the early instrumental era. The stations are: Cadiz/San Fernando, southern Spain 1786–2000; Padova, northern Italy 1725–1997; Milan, northern Italy 1763–1998; Central Belgium, a composite covering 1767–1998 based on several nearby stations; Uppsala, central Sweden 1722–2000; Stockholm, central Sweden 1756–1998; and St Petersburg, Russia 1743–1996. The dataset contains homogenized and quality controlled daily mean, maximum and minimum temperatures (for Stockholm, Uppsala and St Petersburg only daily mean temperature is available). These long-term records are used here to put the model simulated climate change signals in the perspective of longer-term climate variability.

2.3 Linking model gridcell data to station data

For each ECA station, the covering gridcell of each model was used for the bias assessment. Point measurements, however, may not be comparable to the size of a model grid box, which is about 50×50 km in the RCMs studied here as the models do not simulate subgrid-scale variability. There is a particular problem in coastal regions where observational stations are influenced both by land and sea. The model grid box covering the coastal station could be either land or sea depending on the land–sea mask in the model. We have therefore excluded all combinations of coastal stations and model sea grid boxes from our comparisons. For models with fractional land/sea gridboxes we excluded gridcells with more than 50% sea.

Regions of complex topography also give rise to various problems. Differences between model and station altitude have been accounted for using a lapse rate of 0.0065 K/m. This average lapse rate does not account for other local deviations or specific weather conditions. As we are focusing on comparisons of climatological differences and not per-observation differences, this simplification is reasonable with the exception of where there is a large altitudinal difference between the station and model gridcell. We therefore exclude model gridcells where this difference exceeds 1,000 m. Further, high altitude stations, above 2,000 m, have been excluded from the comparisons since these stations often tend to sample free tropospheric conditions or are strongly affected by snow cover and glaciers.

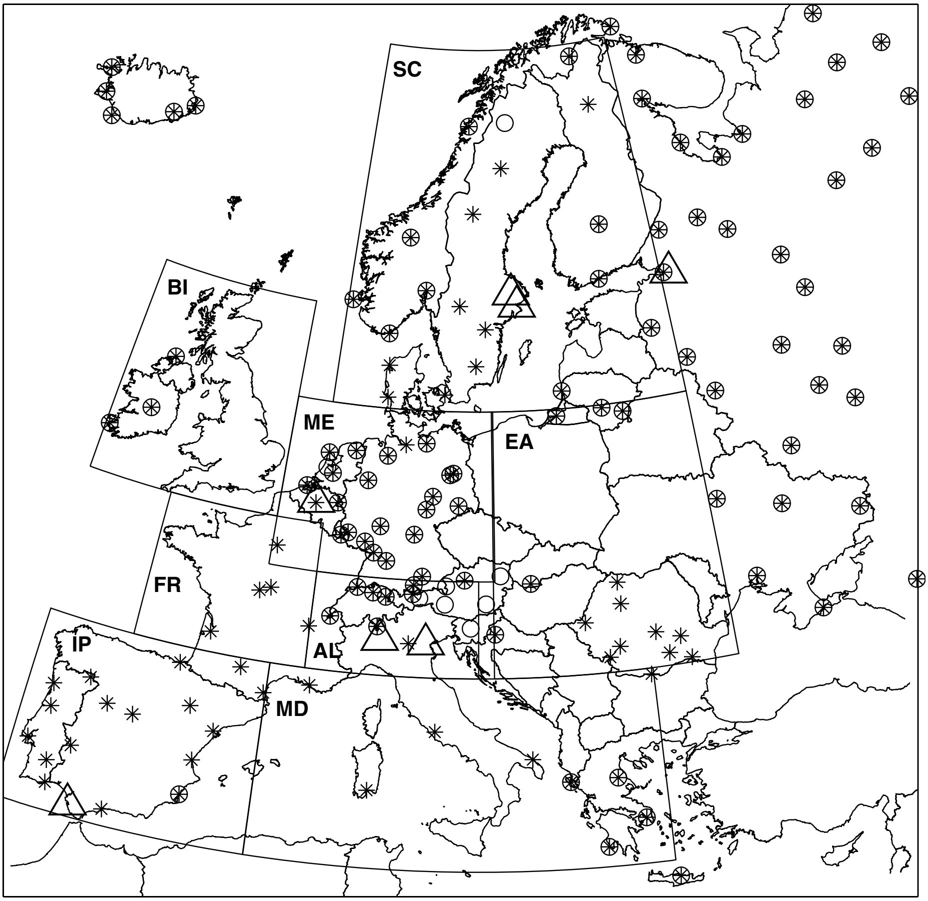

Fig. 1 Location of observational stations in Europe used in this study. *A+* indicates a station where $T_{2\min}$ is available in the ECA data set, an × indicates $T_{2\max}$, and an *O* indicates that $T_{2\mathrm{m}}$ is available. The IMPROVE stations are indicated with *open triangles*. The European subregions are indicated with *letters*: BI – British Isles; SC – Scandinavia; FR – France; EA – Eastern Europe; ME – Mid Europe; AL – Alps; IP – Iberian Peninsula; MD – Mediterranean

These selection criteria result in some variations of the number of models that are used in the ensemble averages. Thus, a final criterion for including any station is that at least 70% of the models are included in the ensemble average for that station.

In the assessment of model behaviour we focus on the median and inter-quartile range (IQR) as robust measures of the model ensemble average (typical) bias and inter-model variation (see Ferro et al. 2005). In this way we avoid undue influence from any one deviating model. The spatial patterns of the median maps and corresponding mean maps (not shown) are however very similar. This is also the case for the IQR and standard deviation maps (not shown), although the IQR maps show a wider inter-model variation compared to the corresponding standard deviation maps. To further reduce the possible influence from station specific local conditions we average, again using medians, both the model bias and simulated climate change on a regional basis. The regions are those used by Déqué et al. (2007) and Jacob et al. (2007), with a slight relaxation of the region boundaries

(Fig. 1) to allow the inclusion of three ECA stations that otherwise were just outside any region. For the British Isles and France regions there are only few stations available, and in the Mediterranean region stations are located along the coast, apart from the interior of the Iberian Peninsula.

We do not fit any probability density function to the observed and simulated temperature distributions, but rather choose to present the simulated empirical distributions in terms of selected percentiles. We have analyzed the 1, 5, 10, 25, 50, 75, 90, 95 and 99 percentiles but choose to focus our presentation on conditions at the wintertime (DJF) fifth and summertime (JJA) 95th percentiles. In this way we can describe features not just of the most extreme situations, but more generally of events occurring on average 4–5 days per season (the 3-month periods DJF and JJA).

3 Results

3.1 Present-day climate, daily maximum temperature in summer

In summer the 95th percentiles for simulated $T_{2\max}$ are highest in southern Europe with maximum temperatures of about or above 40°C (Fig. 3a). The general spatial patterns are reproduced by the RCMs though there are some large differences, locally up to 10°C (Fig. 2a). HadRM3H has relatively small deviations from its driving model, of a similar level to those in RCAO and smaller than the other RCMs, probably related to the fact that it shares common physical parameterizations. The ensemble median of the 95th percentile of $T_{2\max}$ during JJA underestimates the highest temperatures in northern and northwestern Europe and overestimates the high temperatures in southern and eastern Europe compared to ECA (Fig. 4a and Table 1). For most stations except in the White Sea region, the ensemble median bias is within ±3°C. There is, however, a rather large spread among the different models as summarized for each region in Table 1. The negative bias over the British Isles and Scandinavia is evident in almost all models. The underestimation of maximum temperatures in RCAO for northernmost Europe is discussed by Räisänen et al. (2003). They found that this underestimation is related to too cloudy and rainy conditions in this simulation. In a reanalysis-driven experiment, however, Jones et al. (2004) find that RCA2 (the atmospheric component of RCAO) simulates both cloud cover and precipitation in close agreement to observations in northern Europe, suggesting that the driving global data from HadAM3H contributes to the too low maximum temperatures, an effect that could influence all RCMs. Further, the 95th percentile of $T_{2\max}$ from the reanalysis-driven experiment by Jones et al. (not shown) is generally 1–2°C warmer in Scandinavia and the British Isles than that of the RCAO run driven by HadAM3H used in this study. This supports the idea that the driving boundaries contributes to the underestimation of maximum temperatures. Nevertheless, Lenderink et al. (2007) show that the interannual monthly temperature variability in RACMO2 driven by re-analysis boundary conditions is close to the variability in the HadAM3H driven run. Further inspection of the results of these two runs with RACMO2, revealed no significant differences in the simulated $T_{2\max}$ 95th percentile over Europe except for the southeastern part.

In the Iberian Peninsula, the Mediterranean region and eastern Europe the tendency is reversed with an overestimation of the $T_{2\max}$ 95th percentiles in most models. This is particularly the case for the HadRM3H model that is also (much) too warm in the three intermediate regions; France, Mid-Europe and the Alps. These regions are interesting since

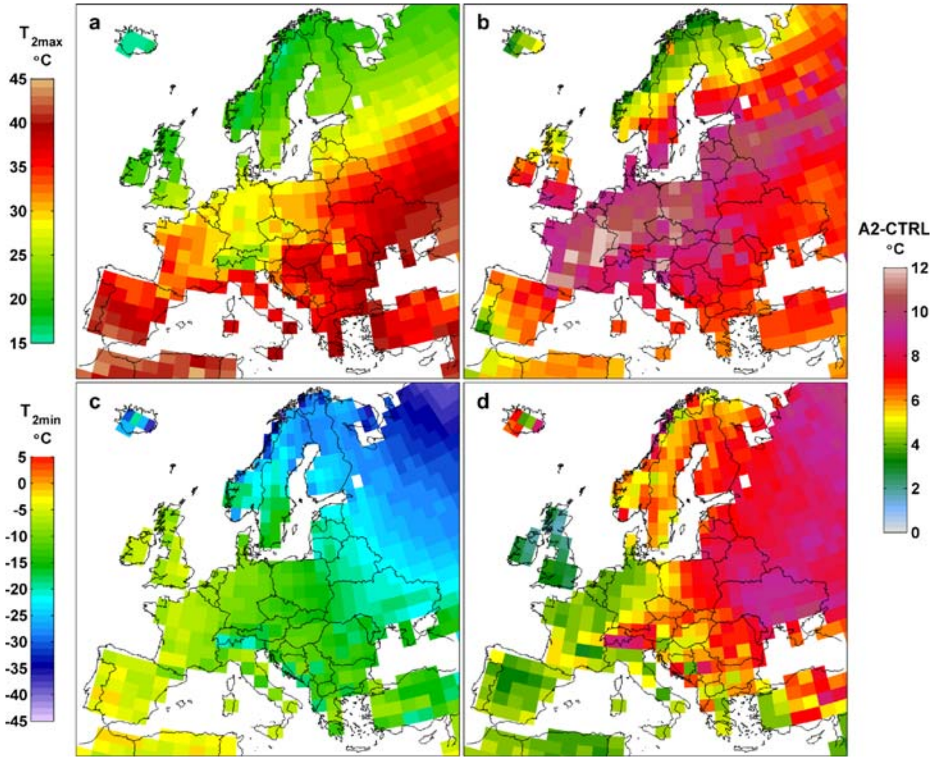

Fig. 2 **a** 95th percentile of T_{2max} in summer (JJA); **b** change in the 95th percentile of T_{2max} in summer (JJA); **c** fifth percentile of T_{2min} in winter (DJF); **d** change in the fifth percentile of T_{2min} in winter (DJF). All panels show results from HadAM3H. Unit: °C

some models show a warm bias (HadRM3H, RegCM2 and RCAO) while some show a cold bias (CLM, PROMES, HIRHAM-NO and RACMO2). The warm bias in RCAO may partly be related to the relatively small storage capacity of the soils in that model (van den Hurk et al. 2004) and, partly to the clear sky portion of the RCA2 solar radiation code that transmits too much radiation to the surface (Jones et al. 2004). Furthermore, Moberg and Jones (2004) noted that a warm bias in southern Europe in HadRM3P (which is very similar to HadRM3H) is associated to low or zero simulated soil moisture. Lenderink et al. (2007) investigate interannual variability of monthly mean temperature in the PRUDENCE simulations. They conclude that on the one hand soil drying and evaporation, and on the other hand clouds and their radiative properties, are major issues in determining the interannual variability of temperature in these simulations. They find that CLM and PROMES have a strong cloud (radiative) control and that RACMO2 has a strong evaporative control on temperature variability. It is therefore expected that these models are colder than the others in these intermediate regions. Our results indicate that their findings about interannual monthly mean temperature variability can be extended also to the more extreme daily temperature conditions discussed in the present study. It is interesting to note that the two versions of HIRHAM show very different patterns of bias magnitude. We have not scrutinized the two simulations in detail. With the exception of the HadRM3H model, all models have an overall median bias within ±2°C. This is an interesting result as previously it was noted that HadRM3H follows most closely its driving model. Thus the

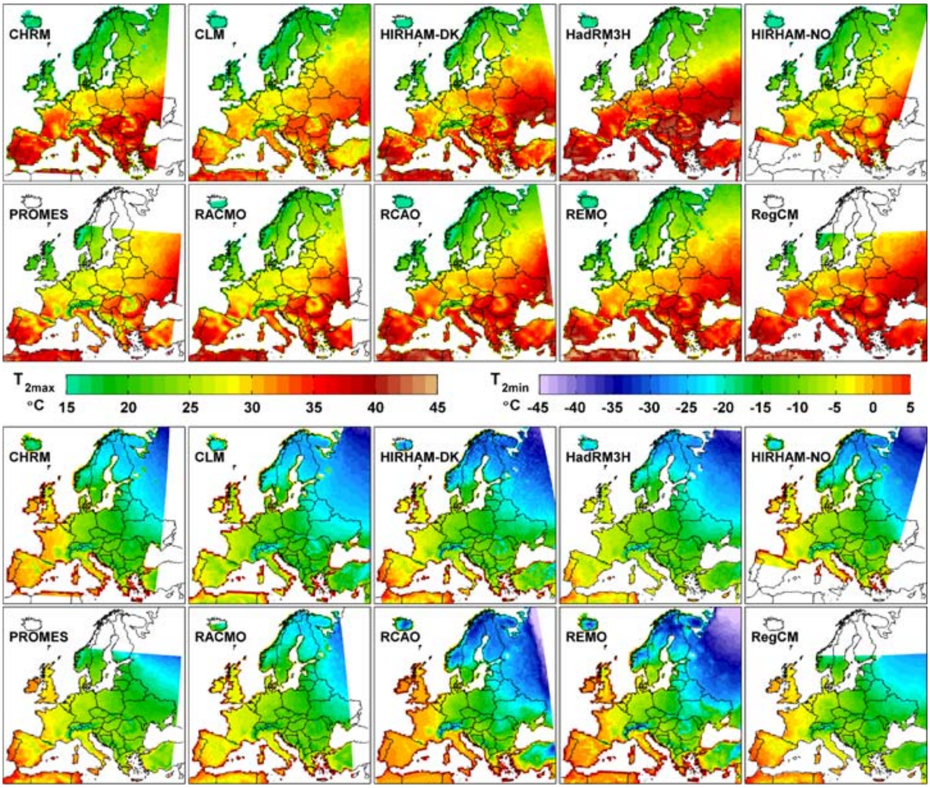

Fig. 3 *Upper panels* show 95th percentiles of $T_{2\max}$ in summer (JJA), and *lower panels* show fifth percentiles of $T_{2\min}$ in winter (DJF) for the 10 RCMs. Unit: °C

other models are deviating more from their driving model but are simulating results which are closer to observation. It would appear that there is a benefit from having the same formulation of the physics in the RCM as in the GCM to improve consistency. However, it still leaves the question open as to whether, in this case, the RCM concerned provides a more realistic projection of regional climate change given its larger biases whilst simulating the current climate.

The ensemble median bias across models and stations (i.e. the median of all the models' biases) is 0.3°C. The overall spatial pattern of biases in the 95th percentile for $T_{2\max}$ is similar to that of the median bias in $T_{2\max}$ (not shown). A difference, however, is that the warm bias in the southeast is stronger and that it extends further to the north in eastern Europe/Russia in the 95th percentile. The fact that model biases affect large parts of the probability distribution is shown in Fig. 5a where regional biases are shown for nine different percentiles. In eastern Europe the positive bias for high percentiles is large, whereas it is small or negative for low percentiles. In several regions, the spread among the RCMs is larger at the 95th and 99th percentile as compared to the median and low percentiles. The cold bias for PROMES and CLM in some regions, as noticed at the 95th percentile, is even more pronounced at the lower temperature percentiles. HadRM3H shows a very pronounced increase in bias for the higher percentiles in most regions. The latter is most likely related to the drying out of soils (see Lenderink et al. 2007; Vidale et al. 2007).

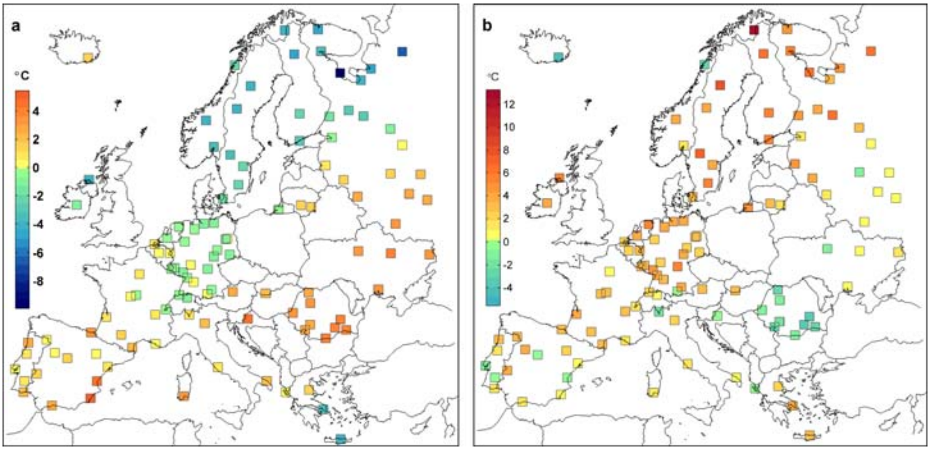

Fig. 4 Ensemble median bias (CTRL-ECA) in **a** the 95th percentile of T_{2max} during summer (JJA) and **b** the fifth percentile of T_{2min} during winter (DJF). Unit: °C

3.2 Present-day climate, daily minimum temperature in winter

The fifth percentiles for simulated T_{2min} in winter are lowest in eastern and northern Europe with temperatures well below −30°C over parts of Russia in most models (Fig. 3). In the western maritime climate the RCMs give fifth percentiles of minimum temperatures close to 0°C. The main features of the geographical distribution of the fifth percentiles of minimum temperatures are similar to the HadAM3H simulation (Fig. 2c) though again there are local differences of up to 10°C. As in summer, HadRM3H deviates little from its global

Table 1 Regional medians of bias (CTRL-ECA station) in each model experiment for the summer (JJA) T_{2max} 95% percentile

N stations	BI	IP	FR	ME	SC	AL	MD	EA	Median	IQR	Range
	3	17	4	25	19	6	9	13			
CHRM	−4.7	1.6	1.9	−0.3	−4.9	0.3	2.3	4.1	0.9	4.6	9.0
CLM	−4.9	−1.4	−1.2	−0.7	−4.7	−2.3	−1.3	0.8	−1.3	2.6	5.7
HIRHAM-DK	−0.2	1.4	0.2	0.4	−1.3	−0.9	3.6	3.4	0.3	3.0	4.8
HadRM3H	−1.3	4.5	4.8	3.3	−1.9	3.6	6.0	7.2	4.1	4.5	9.0
HIRHAM-NO	−4.1	−0.7	−1.3	−2.1	−3.5	−3.3	0.5	−0.1	−1.7	3.0	4.6
PROMES	−2.7	0.4	−1.4	−2.2	–[a]	−1.5	0.8	0.4	−1.4	2.4	3.5
RACMO2	−1.0	0.3	−0.8	−1.7	−2.3	0.0	1.5	0.9	−0.4	1.9	3.8
RCAO	−4.3	2.2	1.7	0.3	−3.6	0.8	0.8	3.0	0.8	3.6	7.3
REMO	−2.4	1.8	0.3	0.1	−3.3	1.2	−0.1	4.6	0.2	2.8	7.9
RegCM	0.0	2.0	2.3	0.4	–[a]	1.9	4.9	3.7	2.0	2.5	4.8
Median	−2.6	1.5	0.3	−0.1	−3.4	0.1	1.2	3.2	0.3		
IQR	3.3	1.7	3.1	2.1	2.1	2.7	3.1	3.3		3.2	
Range	4.9	5.9	6.2	5.5	3.6	6.9	7.3	7.3			12.1

The median, interquartile range (IQR) and total range taken from the individual medians are given per model (right) and per region (bottom), as well as the overall values across all models and regions (lower right corner). The maximum number of ECA stations used in each region is given in the second row.

[a] The simulations with PROMES and RegCM do not include all of Scandinavia.

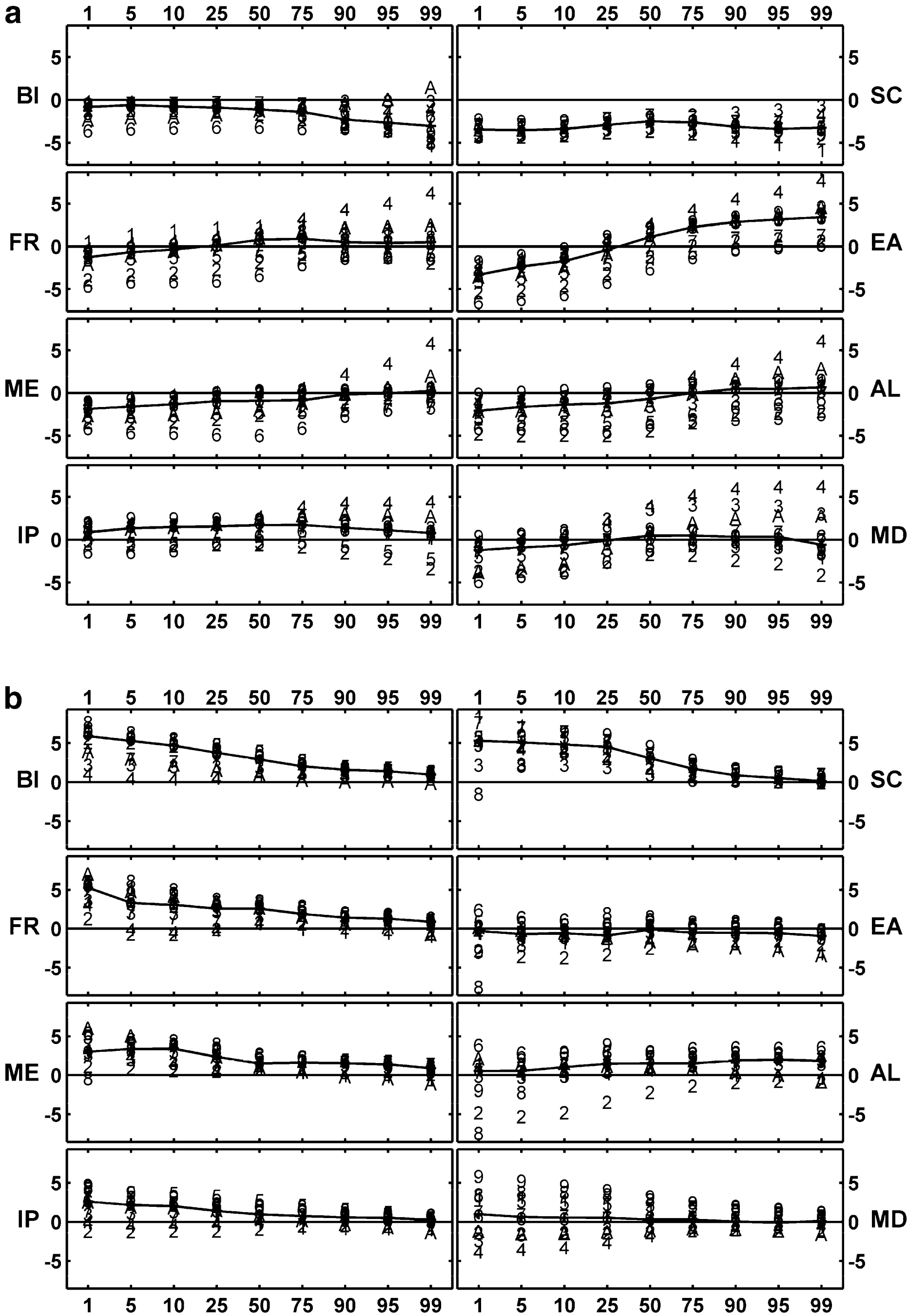

Fig. 5 For each model, the median bias (CTRL-ECA) at different percentiles in **a** $T_{2\max}$ during summer (JJA) and **b** $T_{2\min}$ in winter (DJF) in the eight subregions (cf. Fig. 1). The models are marked as follows: 1-CHRM; 2-CLM; 3-HIRHAM-DK; 4-HadRM3H; 5-HIRHAM-NO; 6-PROMES; 7-RACMO2; 8-RCAO; 9-REMO; A-RegCM2. The *black curve* shows the median bias across the ten-model ensemble. Unit: °C

driving model and is in this case closer than the other RCMs. Cold temperatures in winter are often associated with situations with snow-covered ground and clear skies. A look at the maximum snow depth in the models reveals some differences (not shown). In some models, most notably RegCM2, CHRM, RCAO, and REMO, snow is very rare in low-lying parts of the British Isles, Western France, Spain and Italy, while in others, like CLM, PROMES, RACMO2 and HIRHAM-DK snow occurs more often in these regions. Another process that strongly affects the occurrence of cold winter temperatures is the freezing of soil moisture (Viterbo et al. 1999). There are substantial model-to-model differences in the representation of this process. In Fig. 4b we compare the ensemble median of the DJF fifth percentiles of T_{2min} to the ECA observations. The pattern of biases is similar to that in summer, though reversed in the sense that the bias in western and northern Europe is now positive and the average bias in the southeast is negative. Typical median biases for most stations are in the range ±3°C, apart from generally higher values in Scandinavia.

As in summer for the 95th percentile of T_{2max}, the inter-model variability in the fifth percentile for T_{2min} in winter is large (Table 2), particularly for many coastal stations (not shown). The RCMs generally agree on the sign of the bias in most regions except for the Alps, Mediterranean region and Iberian Peninsula. Notably, in the Alps CLM and RCAO show a marked opposite (i.e. negative) sign of biases compared to the other models, and in the Iberian Peninsula, CLM and HadRM3H show a negative bias while the other models have positive biases. In the Mediterranean and eastern Europe the bias pattern is more diverse: with PROMES and RACMO2 showing little bias in both regions; CHRM, HIRHAM-NO, RCAO and REMO having a positive bias in the Mediterranean region and negative bias in eastern Europe; and CLM, HIRHAM-DK, HadRM3H and RegCM having a negative bias in both regions. For RCAO, the too mild climate in the north can be explained by the circulation being too strongly influenced by westerly winds during winter (cf. Räisänen et al. 2003). Since the forcing conditions from the global model are the same in all experiments this would imply that some fraction of these warm biases is an effect of the boundary conditions given by the driving global model. Van Ulden et al. (2007) show that the HadAM3H simulation is characterized by a too strong westerly flow in winter.

Table 2 Same as Table 1 but for winter (DJF) T_{2min} 5% percentiles

N stations	BI 3	IP 17	FR 4	ME 25	SC 19	AL 6	MD 9	EA 13	Median	IQR	Range
CHRM	5.9	3.0	5.3	2.9	6.0	0.3	4.7	−2.2	3.8	4.0	8.2
CLM	6.3	−2.3	−0.7	0.8	3.3	−4.0	−2.8	−4.3	−1.5	5.5	10.6
HIRHAM-DK	2.2	1.3	2.3	2.0	4.0	−0.2	−0.9	−0.2	1.7	2.7	5.4
HadRM3H	0.4	−1.0	0.1	2.2	3.9	0.6	−2.9	−1.7	0.3	2.8	6.8
HIRHAM-NO	5.8	2.8	3.4	3.7	5.8	1.3	2.6	−0.4	3.1	2.8	6.2
PROMES	5.5	2.4	3.2	5.0	–[a]	3.2	0.4	0.2	3.2	3.6	5.3
RACMO2	1.9	0.2	2.4	3.5	5.9	1.0	0.4	−0.8	1.4	2.6	6.7
RCAO	5.7	3.8	6.1	2.7	2.5	−4.0	4.8	−3.1	3.2	5.6	10.2
REMO	5.0	4.1	4.8	4.3	6.9	1.6	5.3	−1.3	4.5	2.3	8.2
RegCM	2.9	1.9	4.7	5.1	–[a]	0.8	−1.2	−2.2	1.9	4.9	7.3
Median	5.3	2.2	3.3	3.2	4.9	0.7	0.4	−1.5	2.3		
IQR	3.6	2.7	2.7	2.1	2.4	1.5	5.9	1.5		4.2	
Range	5.9	6.4	6.8	4.3	4.3	7.2	8.2	4.5			11.2

[a] The simulations with PROMES and RegCM do not include all of Scandinavia.

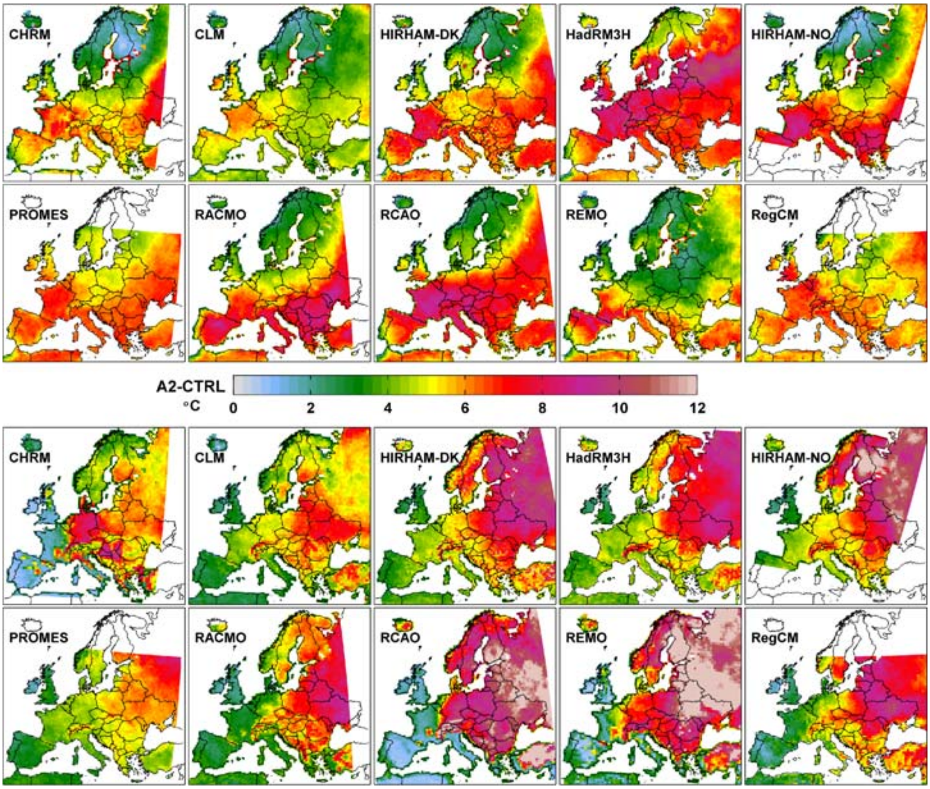

Fig. 6 Change between 1961–1990 and 2071–2100 (SRES A2-CTRL) in **a** the 95th percentile of T_{2max} in summer (JJA) and in **b** the fifth percentile of T_{2min} in winter (DJF) in the 10 RCMs. Unit: °C

They estimate that the bias in circulation induces a positive mean temperature bias of 1.5°C in central Europe. In agreement, using the RACMO2 model central Europe and Southern Scandinavia are about 1.5°C warmer in the presently discussed HadAM3H driven simulation compared to a simulation employing the ERA40 re-analysis for boundary conditions (van Ulden et al. 2006). The different sizes of biases in different models show that the RCMs have different sensitivity to the boundary conditions. Compared to biases in mean conditions the biases in the fifth percentile are larger in most of the regions (Fig. 5b). Again, the spread among models is largest at the more extreme part of the low tail of the distribution with differences at the first percentile ranging up to about 10°C in some areas.

3.3 Future climate change scenario

The projected warming in summer is most dramatic in central and southern Europe in the RCMs. The model range of the change in mean temperature for JJA (not shown) is between 4 and at least 7°C in the Iberian Peninsula, France, the Mediterranean region and in the Alps (Déqué et al. 2007). In T_{2max} the mean signal (not shown) is even larger and for the 95th percentile of T_{2max} it reaches 10°C in a few areas in some models (Fig. 6). Figure 6 reveals large differences between the different RCMs. Both the overall pattern and the size of the changes differ notably. A common feature is that all models simulate a generally larger increase in T_{2max} on warm days in southern Europe than in the north, with the exception of HadRM3H. However, it is again interesting to note that for the other RCMs this is a deviation from the driving GCM; the simulated climate change signal in HadRM3H is very similar to the results obtained with HadAM3H. Most RCMs show a band-like structure of most pronounced warming that extends from France eastwards over the continent, although this structure is more or less displaced to the south. The differences between the ten RCMs are actually as large as the differences between the four different realizations of climate change with one of the RCMs (RCAO) driven by two different emission scenarios (A2 and B2) and lateral boundary conditions from two GCMs (Kjellström 2004). Resulting differences between the RCM with the lowest (CLM) and highest (HadRM3H) T_{2max} changes are about 10°C in large parts of Europe. Just as for the bias in the control climate simulations there is a clear tendency for the climate change signal in summer T_{2max} to be larger for the more extreme situations (95th and 99th percentiles) than in the median (Fig. 7a). From the current analysis it is not obvious whether this is due to variability increases on the interannual or daily range, or a combination of the two. The spread among the models is larger in the upper tail of the distribution compared to the lower tail for all regions. The largest inter-model differences occur in central Europe. This is consistent with the discussion on the role of parameterized physical processes in Section 3.1.

In winter all models show substantial increases in mean temperatures over eastern Europe and Scandinavia. In T_{2min} the mean warming signal is above 4°C in large parts of eastern Europe and Russia (Déqué et al. 2007). Just like in summer the change in the more extreme conditions is larger compared to the change in the average. The changes for the fifth percentile are shown in Fig. 6. The largest differences compared to the control climate are very large, up to 15°C in some locations in REMO and RCAO. At the same time it can be seen that the spread between the models is again very large with differences between models of up to 10°C. As in summer, HadRM3H is close to HadAM3H. The considerable spread between the RCMs is about twice as large as the spread between the four experiments with one only RCM forced by different emission scenarios and different lateral boundary conditions in Kjellström (2004). Common to all the models is a connection between the region of maximum climate change signal and the withdrawal of the snow cover in the models. This connection indicates the importance of the feedback processes involving temperature, snow cover and albedo on the temperature climate in these models. Decreasing snow cover leads to lower albedo which allows more shortwave radiation to be absorbed at the ground. Also, the reduced snow cover facilitates heat exchange between the relatively warm soil and the atmosphere. Both these effects lead to higher temperatures which in turn act to further reduce the snow cover. The importance of these feedback processes is illustrated by the fact (not shown graphically) that the areas of maximum changes in the fifth percentile of T_{2min} are, in most models, covered with snow during more than 50% of the time in CTRL. In A2 these regions generally have snow in less than 25% of the time. Again, the climate change signal is stronger and the spread between the models larger near the lower tail of the distribution (Fig. 7b).

3.4 Future climate change in relation to historical changes

The observed temporal variability between different 30-year periods at the seven IMPROVE stations is compared with the inter-model variability in Fig. 8. For comparison with observed natural variability we use as reference level the median of the percentiles

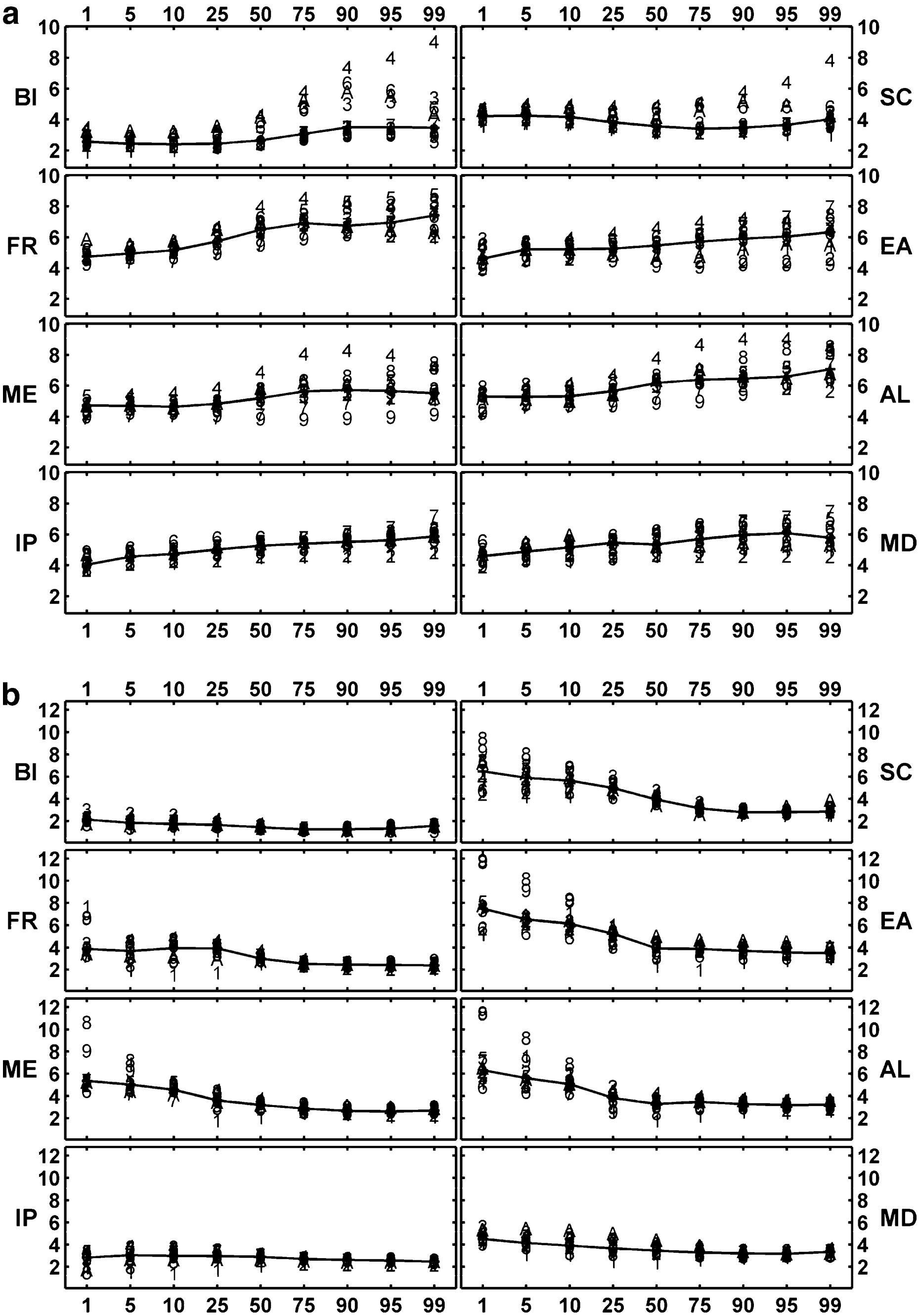

Fig. 7 For each model, the average change between 1961–1990 and 2071–2100 (SRES A2-CTRL) in the eight subregions (cf. Fig. 1). The models are marked as in Fig. 5. The *black curve* shows the median change across the 10-model ensemble. Unit: °C

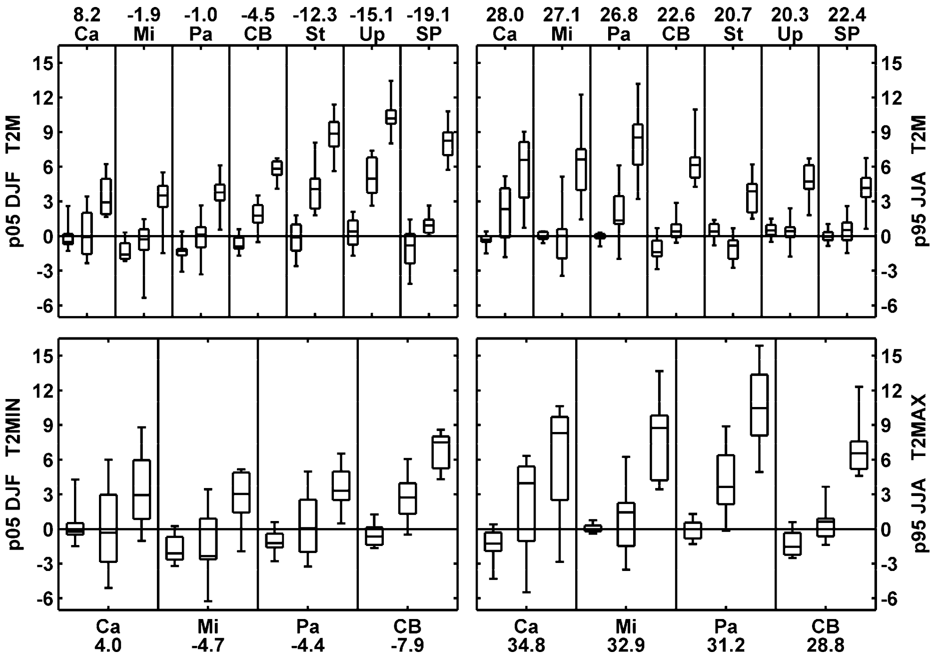

Fig. 8 Simplified boxplots of the 5th percentiles of the winter daily mean temperature, T_{2m} (*upper left*) and daily minimum temperature (*lower left*), and the 95th percentile of the summer daily mean (*upper right*) and daily maximum (*lower right*) temperature. In each section of the panels *three boxplots* for a station is shown. The stations are Cadiz (Ca), Milan (Mi), Padua (Pa), Central Belgium (CB), Stockholm (St), Uppsala (Up) and Saint Petersburg (SP). For each station, the *left boxplot shows* the observed spread between different overlapping 30-year periods. The first period is 1751–1780; the second period is 1761–1790, etc. up to 1971–2000. As indicated in the text, for the first two and the last periods data is lacking for some stations. The *middle boxplot* shows the spread between the different RCMs for the same period (CTRL), and to the right is the RCM spread for the future SRES A2 simulations. On the *y*-axis is the deviation from the observed 1961–1990 median which is shown with the station acronym. In each boxplot the box extends from the lower to the upper quartile, with the *line inside the box* denoting the median. *Lines* indicating the tails extend outward from the quartiles to the minimum/maximum value. Unit: °C

observed for the period 1961–1990. All box plots are presented as deviations from these observed median temperatures for 1961–1990. The observed climate variability, including the observed climate change during the typically 200 year long observational period is substantially smaller than the inter-model variability. This is particularly true for the Mediterranean stations. The relatively long positive tail for Cadiz in winter is an effect of the first decades of observations in that series being very warm, for the rest of the observational period the variability is significantly smaller. The observed temporal variability of the 5th or 95th percentiles of T_{2m}, expressed as IQR between 30-year periods, is typically 1–2 K, which can be compared to 2–4 K for the simulated inter-model variability (IQR). For T_{2min} (5th percentile) and T_{2max} (95th percentile) the simulated inter-model variability (IQR) is even larger, 3–6 K. The corresponding variability of the median (not shown) is less pronounced. The observed temporal IQR of the median is about 1 K or less, and the inter-model IQR is about 2 K for T_{2m} and about 3 K for T_{2min} and T_{2max}. This difference is consistently seen when comparing the IQR of the median and the 95th/5th percentile at a station. Despite the substantial inter-model variability, the climate change signal (A2-CTRL) is well beyond the natural variability. With the exception of Cadiz/San Fernando, the simulated change is more pronounced in the 5th and 95th percentiles than in the median (not shown). The inter-model median climate change is well beyond both the observed temporal IQR and the inter-model IQR. For Cadiz/San Fernando the comparatively weak climate change signal and large modeled IQR is likely an effect of the maritime location of this station, with the strong dampening effect of the Atlantic. In the northeast, at Stockholm, Uppsala and Saint Petersburg, where only T_{2m} is available, the simulated climate change signal is substantial. Even though the IMPROVE stations are not necessarily representative for the regions in Fig. 1, they do follow the general pattern of regional biases (Tables 1 and 2).

4 Summary and conclusions

The simulated daily maximum and minimum temperatures in ten RCMs are compared to the observed climate for the time period 1961–1990. It is found that the models generally underestimate (overestimate) the maximum temperatures in northern (southern) Europe during summer. In winter, minimum temperatures are overestimated in large parts of western and northern Europe while there is an underestimation in the southeast. It is also found that the biases are larger in the 95th/5th percentiles than the corresponding biases in the median, i.e. the biases generally increase towards the tails of the probability distributions. We also show that there are large inter-model differences of, locally, up to ±10°C in the 95th/5th percentiles. Despite these large inter-model differences we find that the biases in the simulated 95th/5th percentiles are smaller than ±3 (±5°C) when averaged over most (all) European sub-regions. The regional biases in 95th/5th percentiles are significantly smaller than the biases in the absolute maxima (i.e. highest recorded) of daily maximum temperatures during summer. The latter may be as large as 10°C as reported for the HadRM3P model by Moberg and Jones (2004) for some single stations in southern and southeastern Europe. We note that the HadRM3H model analysed here (which is very similar to HadRM3P) also has exceptionally warm summer T_{2max} biases in southern and southeastern Europe, much larger than any of the other nine models considered. Hence, the problem with positively biased summer temperatures discussed by Moberg and Jones (2004) is not as serious for the other models.

We also investigate the climate change signal in the PRUDENCE common experiment with the ten RCMs. The RCMs simulate considerable changes in extreme temperatures both in summer and winter, and in both cases these are larger than changes in the median for large areas. As for the biases we also find large inter-model differences. These differences are briefly compared to a previous study including four experiments with one RCM driven by two different global models and two emission scenarios The inter-model differences in T_{2max} and T_{2min}, at the 95th and 5th percentiles, respectively, are as large as the differences between these aforementioned experiments. This implies that, for the extreme quantiles, the uncertainty in the amplitude of the climate change signal due to regional climate model formulation is as large as the uncertainty due to emission scenario (A2 and B2) or GCM boundary forcing. Differences between models results are amplified at the extremes in both simulation of present-day climate and in projections of future climate change.

A comparison with historical data shows that the spread of the simulated extreme temperatures is larger than the natural variability during the last centuries, at least for the observational stations with long enough records. Nevertheless, the simulated future climate change signal is found to be well beyond the natural variability at these locations.

Acknowledgements Parts of this work were undertaken in the European PRUDENCE project (project EVK2-CT2001-00132 in the EU fifth Framework program for Energy, environment and sustainable environment and the European MICE project (EKV2OCT2001-00018). Supplemental funding from the Swiss National Science Foundation (NCCR Climate) and from the UK government's Department for Environment, Food and Rural Affairs (Defra) Climate Prediction Programme (PECD 7/12/37) is also acknowledged. We are grateful to all our colleagues in the PRUDENCE project that has provided the data used in this work. In particular Ole Christensen at DMI has been very helpful in organizing and maintaining the PRUDENCE data distribution centre (<prudence.dmi.dk>). Anders Moberg, Dave Rowell, Burkhardt Rockel, Markku Rummukainen and Bart van den Hurk are acknowledged helpful comments. The authors are indebted to two anonymous reviewers for their most valuable comments on the manuscript.

References

Beniston M, Stephenson D, Christensen O, Ferro C, Frei C, Goyette S, Halsnaes K, Holt T, Jylhä K, Koffi B, Palutikof J, Schöll R, Semmler T, Woth K (2007) 'Current and future extreme climatic events in Europe: observation and modeling studies conducted within the EU "PRUDENCE" project.' Clim Change, doi:10.1007/s10584-006-9226-z (this issue)

Buonomo E, Jones RG, Huntingford C, Hannaford J (2007) The robustness of high resolution predictions of changes in extreme precipitation for Europe. Q J R Meteorol Soc (in press)

Camuffo D, Jones P (2002) Improved understanding of past climatic variability from early daily European instrumental sources. Clim Change 53(1–3):1–4

Christensen J, Carter TR, Rummukainen M (2007) Evaluating the performance and utility of regional climate models: the PRUDENCE project. Clim Change, doi:10.1007/s10584-006-9211-6 (this issue)

Déqué M, Rowell D, Schär C, Giorgi F, Christensen JH, Rockel B, Jacob D, Kjellström E, de Castro M, van den Hurk B (2007) An intercomparison of regional climate models for Europe: assessing uncertainties in model projections. Clim Change, doi:10.1007/s10584-006-9228-x (this issue)

Ferro CAT, Hannachi A, Stephenson DB (2005) Simple non-parametric techniques for exploring changing probability distributions of weather. J Clim 18:4344–4354

Giorgi F, Hewitson B, Christensen J, Hulme M, von Storch H, Whetton P, Jones R, Mearns L, Fu C (2001) Regional climate information – evaluation and projections. In: Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai K, Maskell K, Johnson CA (eds) Climate change 2001: the scientific basis. contribution of working group I to the third assessment report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK p 881

Hanssen-Bauer I, Førland E, Haugen JE, Tveito OE (2003) Temperature and precipitation scenarios for Norway: comparison of results from dynamical and empirical downscaling. Clim Res 25:15–27

Jacob D, Bärring L, Christensen OB, Christensen JH, de Castro M, Déqué M, Giorgi F, Hagemann S, Hirschi M, Jones R, Kjellström E, Lenderink G, Rockel B, Sánchez E, Schär C, Seneviratne SI, Somot S, van Ulden A, van den Hurk B (2007) An intercomparison of regional climate models for Europe: design of the experiments and model performance. Clim Change, doi:10.1007/s10584-006-9213-4 (this issue)

Jones CG, Wyser K, Ullerstig A, Willén U (2004) The Rossby Centre regional atmospheric climate model (RCA). Part II: application to the Arctic climate. Ambio 33(4–5):211–220

Kjellström E (2004) Recent and future signatures of climate change in Europe. Ambio 33:193–298

Klein Tank AMG, Wijngaard JB, Können GP, Böhm R, Demarée G, Gocheva A, Mileta M, Pashiardis S, Hejkrlik L, Kern-Hansen C, Heino R, Bessemoulin P, Müller-Westermeier G, Tzanakou M, Szalai S, Pálsdóttir T, Fitzgerlad D, Rubin S, Capaldo M, Maugeri M, Leitass A, Bukantis A, Aberfeld R, van Engelen AFV, Førland E, Mietus M, Coelho F, Mares C, Razuvaev V, Nieplova E, Cegnar T, Antonio López J, Dahlström B, Moberg A, Kirchhofer W, Ceylan A, Pachaliuk O, Alexander LV, Petrovic P (2002) Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. Int J Climatol 22:1441–1453

Klein Tank AMG, Können GP (2003) Trends in daily temperature and precipitation extremes in Europe, 1946–99. J Clim 16:3665–3680

Lenderink G, van Ulden A, van den Hurk B, van Meijngaard E (2007) Summertime inter-annual temperature variability in an ensemble of regional model simulations: analysis of the surface energy budget. Clim Change (this issue)

Moberg A, Jones P (2004) Regional climate model simulations of daily maximum and minimum near-surface temperatures across Europe compared with observed station data 1961–1990. Clim Dyn 23:695–715

Nakićenović N, Alcamo J, Davis G, de Vries B, Fenhann J, Gaffin S, Gregory K, Grübler A, Jung TY, Kram T, La Rovere EL, Michaelis L, Mori S, Morita T, Pepper W, Pitcher H, Price L, Riahi K, Roehrl A, Rogner H-H, Sankovski A, Schlesinger M, Shukla P, Smith S, Swart R, van Rooijen S, Victor N, Dadi Z (2000) Emission scenarios. A Special Report of Working Group III of the Intergovernmental Panel on Climate Change. Cambridge University Press, pp 599

Noguer M, Jones RG, Murphy JM (1998) Sources of systematic errors in the climatology of a regional climate model over Europe. Clim Dyn 14:691–712

Räisänen J, Hansson U, Ullerstig A, Döscher R, Graham LP, Jones C, Meier M, Samuelsson P, Willén U (2003) GCM driven simulations of recent and future climate with the Rossby Centre coupled atmosphere – Baltic Sea regional climate model RCAO. SMHI Reports Meteorology and Climatology 101, SMHI, SE 60176 Norrköping, Sweden, p 61

Rowell DP (2005) A scenario of European climate change for the late 21st century: seasonal means and interannual variability. Clim Dyn DOI: 10.1007/s00382-005-0068-6

Schär C, Vidale PL, Lüthi D, Frei C, Häberli C, Liniger MA, Appenzeller C (2004) The role of increasing temperature variability for European summer heat waves. Nature 427:332–336. doi: http://dx.doi.org/10.1038/nature02300

van den Hurk B, Hirschi M, Schär C, Lenderink G, van Meijgaard E, van Ulden A, Rockel B, Hagemann S, Graham P, Kjellström E, Jones R (2004) Soil control on runoff response to climate change in regional climate model simulations. J Clim 18:3536–3551

van Ulden A, Lenderink G, van den Hurk B, van Meijgaard E (2007) Circulation statistics and climate change in Central Europe: PRUDENCE simulations and observations. Clim Change, doi:10.1007/s10584-006-9212-5 (this issue)

Vidale PL, Lüthi D, Wegmann R, Schär C (2007) ‘European summer climate variability in a heterogeneous multi-model ensemble. Clim Change, doi:10.1007/s10584-006-9218-z (this issue)

Viterbo P, Beljaars ACM, Mahfouf J-F, Teixeira J (1999) The representation of soil-moisture freezing and its impact on the stable boundary layer. QJR Meteorol Soc 125:2401–2426

Yan Z, Jones PD, Davies TD, Moberg A, Bergström H, Camuffo D, Cocheo C, Maugeri M, Demarée GR, Verhoeve T, Thoen E, Barriendos M, Rodriguez R, Martin-Vide J, Yang C (2002) Trends of extreme temperatures in Europe and China based on daily observations. Clim Change 53:355–392