Climatic Change
DOI 10.1007/s10584-011-0224-4

Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal

Matthias Jakob Themeßl · Andreas Gobiet · Georg Heinrich

Received: 28 January 2011 / Accepted: 9 August 2011

Abstract Realizing the error characteristics of regional climate models (RCMs) and the consequent limitations in their direct utilization in climate change impact research, this study analyzes a quantile-based empirical-statistical error correction method (quantile mapping, QM) for RCMs in the context of climate change. In particular the success of QM in mitigating systematic RCM errors, its ability to generate "new extremes" (values outside the calibration range), and its impact on the climate change signal (CCS) are investigated. In a cross-validation framework based on a RCM control simulation over Europe, QM reduces the bias of daily mean, minimum, and maximum temperature, precipitation amount, and derived indices of extremes by about one order of magnitude and strongly improves the shapes of the related frequency distributions. In addition, a simple extrapolation of the error correction function enables QM to reproduce "new extremes" without deterioration and mostly with improvement of the original RCM quality. QM only moderately modifies the CCS of the corrected parameters. The changes are related to trends in the scenarios and magnitude-dependent error characteristics. Additionally, QM has a large impact on CCSs of non-linearly derived indices of extremes, such as threshold indices.

1 Introduction

Regional climate models (RCMs; Giorgi and Mearns 1991, 1999; Wang et al. 2004) are widely used tools for providing regional climate information over limited areas. With projects as ENSEMBLES (van der Linden and Mitchell 2009) or PRUDENCE (Christensen

Electronic supplementary material The online version of this article (doi:10.1007/s10584-011-0224-4) contains supplementary material, which is available to authorized users.

M. J. Themeßl (✉) · A. Gobiet · G. Heinrich
Wegener Center for Climate and Global Change and Institute for Geophysics, Astrophysics, and Meteorology, University of Graz, Graz, Austria
e-mail: matthias.themessl@uni-graz.at

Published online: 23 September 2011

and Christensen 2007) the availability and reliability of RCM simulations for Europe has increased rapidly in recent years. However, RCMs still feature considerable systematic errors (e.g., Frei et al. 2003; Suklitsch et al. 2008, 2010), which complicate the application of RCM results in climate change impact research.

One common way to deal with model errors in climate change impact studies is the “delta change approach”, also called perturbation method (Déqué 2007; Fowler and Kilsby 2007; Graham et al. 2007). This method generates climate scenarios by adding the climate change signal (CCS) from a RCM simulation to daily or monthly observations. CCS is defined as the difference of climatological means (e.g., monthly, seasonal, or annual) between the future (e.g., 2021–2050) and present or past (e.g., 1971–2000) of a climate variable. By taking the difference, systematic model errors are removed as long as they are similar in both periods, but any potential change in temporal variability is removed as well, since variability is inherited from the observations.

Besides the delta approach, more sophisticated RCM post-processing methods have been proposed and evaluated by e.g., Boé et al. (2007), Graham et al. (2007), Leander and Buishand (2007), Lenderik et al. (2007), Dobler and Ahrens (2008), Piani et al. (2009, 2010), or Themeßl et al. (2011). These approaches belong to the family of Model Output Statistics (MOS; Wilks 1995; Maraun et al. 2010) and are termed “empirical-statistical downscaling and error correction methods” (DECMs). DECMs are technically identical to empirical-statistical downscaling (ESD; Benestad et al. 2008) but relate modeled instead of observed predictors to observations (predictand). As a consequence, DECMs are only valid for the model they are calibrated on and, in addition to the ESD’s traditional purpose of downscaling coarser resolved model results to the local scale, also aim at the reduction of model errors.

In a comprehensive inter-comparison study of seven DECMs for daily precipitation from a 10 km resolved RCM Themeßl et al. (2011) conclude that quantile mapping (QM) outperforms all other investigated DECMs, although local intensity scaling and the analogue methods nearly result in similar performance. Besides, they also show that at least for daily precipitation linear regression approaches, although optimized by predictor transformation and randomization, fail in systematically reducing RCM error characteristics, as expected. Based on these findings, as well as similar shown correction potential for other RCMs and parameters (e.g., Dobler and Ahrens 2008; Piani et al. 2009; Heinrich and Gobiet 2011; or Rojas et al. 2011), QM is chosen here for all error correction purposes.

However, most studies that evaluate DECMs applied on RCMs are based on relatively short simulations of the past. Applications to longer climate simulations also exist but primarily in hydrological literature and rather focus on the results of hydrological models (Dettinger et al. 2004; Wood et al. 2004; Fowler and Kilsby 2007; Leander et al. 2008; van Pelt et al. 2009) than on the performance of the error correction. Thus, the aim of this study is a) to extend the available studies by analyzing the performance of QM in the context of climate change, b) to demonstrate the flexibility of QM to be successfully applied to different parameters c) to investigate QM’s options to reproduce “new extremes” (values outside the calibration range), and d) to analyze the impact of QM on the CCS.

The article is structured as follows: Section 2 focuses on the data used in this study and on methodological issues. Section 3 discusses the performance of the applied DECM for different parameters and its impacts on CCSs, followed by Section 4, which summarizes and discusses the key findings of this study.

2 Data and methods

2.1 Data

Daily mean, minimum, and maximum temperature as well as daily precipitation amount from the COSMO model in climate mode (CCLM, version 2.4.6; Böhm et al. 2006) are used in this study. The model data is provided by ETH Zurich within the ENSEMBLES project and covers entire Europe with a gridspacing of about 25 km. The applied simulations are driven by lateral boundary conditions from the ERA-40 reanalysis (Uppala et al. 2005) for the period 1961–2000 (hindcast simulation) and from the coupled ocean-atmosphere general circulation model (GCM) HadCM3 with normal sensitivity (Q0; Gordon et al. 2000) for the period 1961–2050 (control and scenario simulation). The control simulation (1961–2000) is based on observed greenhouse gas concentrations and the scenario simulation (2001–2050) on the SRES greenhouse gas emission scenario A1B (Nakicenovic et al. 2000). Within the ENSEMBLES project 23 RCM simulations driven by 8 different GCMs are available. Compared to other simulations in this ensemble, the CCLM simulation can be regarded as sensitive in terms of temperature change and moderate in terms of precipitation change (Heinrich and Gobiet 2011).

The E-OBS dataset (version 2; Haylock et al. 2008) provides the observational reference for this study. We use the respective 0.25-degree gridded daily mean, minimum, and maximum temperature as well as daily precipitation amount for the period 1961–2000. The data represents spatially averaged values rather than point-scale information and is therefore suited to be compared or related to RCM results (Déqué 2007). However, E-OBS features known deficiencies, i.e. mean values and temperature parameters are of higher quality than extremes and precipitation amount (Hofstra et al. 2009). As E-OBS lacks some data at the beginning of the 1960s, only grid cells with at least 80% data availability between 1961 and 2000 are used.

For the subsequent evaluation, Europe is divided into eight sub-regions, according to Christensen and Christensen (2007) (Fig. 1).

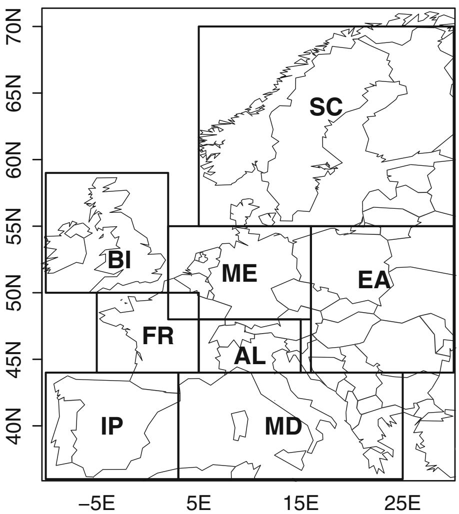

Fig. 1 The location of the European sub-regions. BI (British Islands), IP (Iberian Peninsula), FR (France), ME (Mid-Europe), SC (Scandinavia), AL (Alps), MD (Mediterranean), EA (Eastern Europe)

2.2 Method

2.2.1 Basic Quantile Mapping method

Our implementation of QM can be classified as distribution-based (calibrated on climatological distributions rather than on paired data), direct (predictor and predictand are the same parameters), and parameter-free (using empirical cumulative density distributions, *ecdfs*, rather than theoretical cumulative distribution functions). QM is applied on daily basis (t) and for each grid cell (i) separately resulting in a corrected time series Y^{cor} in Eq. 1 using a correction function (CF) defined in Eq. 2.

$$Y^{cor}_{t,i} = X^{raw}_{t,i} + CF_{t,i} \tag{1}$$

$$CF_{t,i} = ecdf^{obs,cal^{-1}}_{doy,i}\left(P_{t,i}\right) - ecdf^{mod,cal^{-1}}_{doy,i}\left(P_{t,i}\right) \tag{2}$$

$$P_{t,i} = ecdf^{mod,cal}_{doy,i}\left(X^{raw}_{t,i}\right) \tag{3}$$

CF represents the difference between the observed (*obs*) and the modeled (*mod*) inverse *ecdf* ($ecdf^{-1}$) for the respective day of the year (*doy*) in the calibration period (*cal*) at probability P. P is obtained by relating the raw climate model output X^{raw} to the corresponding *ecdf* in the calibration period. For QM calibration, *doy* is centred within a 31 days moving window, which is used to construct an *ecdf* for each day of the year.

2.2.2 Frequency adaptation

This study extends the basic QM procedure described in the previous subsection (QMv0, further details in Themeßl et al. 2011) by frequency adaptation (QMv1). Frequency adaptation (FA) is applied in order to account for a methodological problem, which occurs if the dry-day frequency in the model result ($ecdf^{\,mod,cal}$) is greater than in the observations ($ecdf^{\,obs,cal}$). Usually this is not the case, since many RCMs tend to underestimate the dry-day frequency ("drizzling-effect"; e.g., Gutowski et al. 2003), but occurs, e.g., along with the so called summer drying problem of RCMs in south-eastern Europe (Hagemann et al. 2004; Jacob et al. 2007). In such cases, QM without FA results in a systematic wet precipitation bias as any dry day in X^{raw} is mapped to a precipitation day (compare Fig. 2). With FA only the fraction $\Delta P_0 = \left(ecdf^{mod,cal}_{doy,i}(0) - ecdf^{obs,cal}_{doy,i}(0)\right)/ecdf^{mod,cal}_{doy,i}(0)$ of such dry-day cases with probability P_0 are corrected randomly by linearly interpolating between zero precipitation and the precipitation amount of $ecdf^{obs,cal^{-1}}_{doy,i}\left(ecdf^{mod,cal}_{doy,i}(0)\right)$ (the first precipitation class in QM without FA). By this means, the wet bias is removed (Fig. 2a and b). FA also confirms its applicability and stability in a stricter evaluation setup with differing calibration and evaluation periods (decadal cross validation, see Section 3.1) between 1961 and 2000. Applied to the control RCM simulation and the observational reference E-OBS, the summer season precipitation overestimation in Eastern Europe of QM without FA is removed systematically.

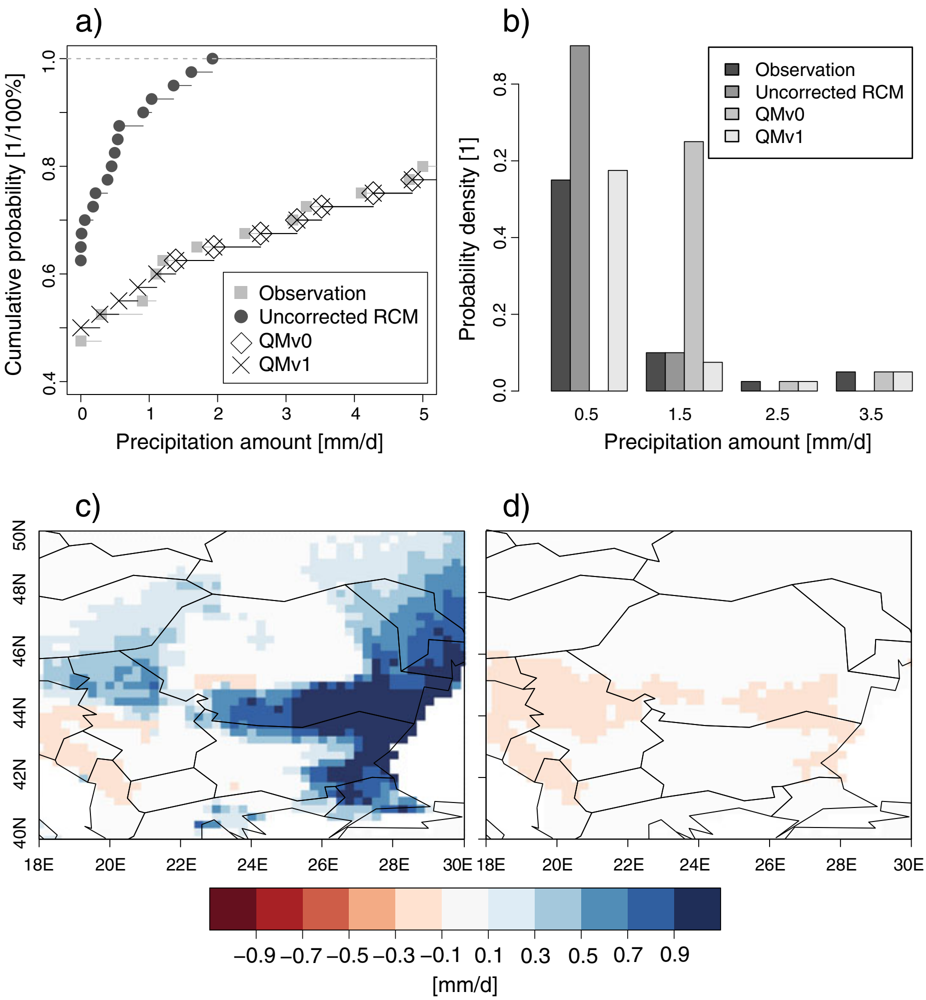

Fig. 2 The impact of FA on QM's performance for daily precipitation. QMv0 and QMv1 are shown. Both QM versions are calibrated and evaluated for July 1 1961–2000 at one single grid-cell in Eastern Europe with pronounced summer drying problem (panels **a** and **b**). Precipitation in panel b is classified into 1 mm/d bins, where the respective class mid is indicated on the x-axis. Panels **c** (QMv0) and **d** (QMv1) show the cross validated summer season precipitation bias maps for Eastern Europe

2.2.3 New extremes

In climate change applications, the question arises how to treat values outside the range of the calibration period ("new extremes"). Unlike for a wide range of values that are not too close to the outer bounds of the calibration range (where QM clearly reduces model errors, see subsequent sections), DECMs are expected to deteriorate RCM results with regard to new extremes, since they are often per construction not able to produce such values. In order to mitigate this problem, two further extensions of QMv1 are proposed here and compared to QMv0. These extensions comprise

a) constant extrapolation of the correction value (difference between $ecdf^{obs,cal}$ and $ecdf^{mod,cal}$) at the highest and lowest quantiles of the calibration range (QMv1a; compare Boé et al. 2007; Déqué 2007) and
b) the same extrapolation, but neglecting the three highest/lowest correction terms (QMv1b). This approach is based on the assumptions that the tails of the correction functions are likely to be very noisy.

Fitting theoretical distributions to the empirical distribution function of the data (e.g., Piani et al. 2009; Dobler and Ahrens 2008) would resolve the "new extremes" problem as well. However, such a fitting is not part of this study, since this would lead to a loss of information compared to empirical distributions and thus potentially introduces biases. Furthermore, by using theoretical distribution QM becomes less flexible in its application to different parameters and regions as a priori information about the shape of the probability density functions (*pdfs)* is needed. Also fitting arbitrary transfer functions to the *ecdf*s resolves the "new extremes" problem (see Piani et al. 2010 for an application to global climate models), but still, problems and shortcomings due to the fitting process have to be taken into account.

The evaluation of QMv1a and QMv1b is based on the ERA-40-driven CCLM hindcast in order to assure highest possible temporal correlation between the model and observations. This is necessary in order to design an evaluation setup that contains "new extremes" in the observational data of the evaluation period, which should be also represented in a well-performing model. The annual maxima of uncorrected, corrected (3 variants), as well as observed daily time series for the 5 years containing the highest annual maxima in the observation between 1961 and 2000 are evaluated at each grid cell over Europe. The evaluation years are always left out in the calibration process, which guarantees that the evaluated extremes are outside of the calibration range. Quantile-quantile (QQ)-plots in Fig. 3 show the results for maximum 1 day precipitation for entire Europe. Comparable results are obtained for the three temperature parameters.

The uncorrected RCM overestimates the maxima in this case, while QMv0 significantly underestimates the new extremes due to its methodological constraint of mapping to the maximum of the calibration period. In contrast, QMv1a and QMv1b are both able to generate new extremes. Furthermore, QMv1a and QMv1b are capable to partly outperform

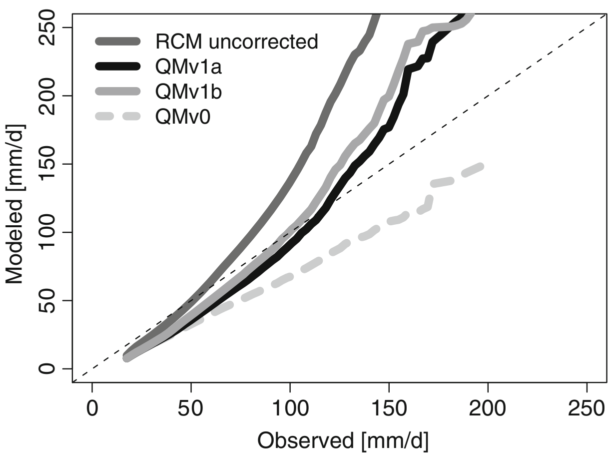

Fig. 3 Comparison of different QM approaches applied to new extremes of daily precipitation amount. The QQ-plot compares yearly maxima of the 5 years with highest maxima for each grid cell in entire Europe between 1961 and 2000

the uncorrected RCM even for new extremes. This is surprising, since the simple extrapolation is based only on weak empirical evidence. However, this finding clearly indicates that both extrapolations are favourable compared to QMv0, since it removes a major drawback of the QM method without deterioration of the RCM's quality. The remaining overestimation in the upper tails is due to the fact, that allowing new extremes in QM is accompanied with the loss of the ability to reliably remove outliers from the RCM output (as done by QMv0).

Comparing QMv1a to QMv1b, QMv1a performs slightly better in this example due to the rather smooth tails of the CCLM error correction functions (not shown). Thus, QMv1a is applied in all our subsequent analyses.

3 Results

3.1 Evaluation of QM with regard to precipitation and temperature

The evaluation focuses on QM's error correction potential, not on downscaling, as the RCM and E-OBS feature comparable spatial resolutions. In order to assess the applicability of QM to a climate simulation setup (i.e. to a GCM-driven RCM), our evaluation compares the control simulation between 1961 and 2000 to the observational reference E-OBS. Based on the climate simulation setup, QM corrects for the combined GCM-RCM error. Such application is only possible with distribution-wise DECMs that do not rely on temporal correlation between the model and the observation. The flexibility of QM is assessed by evaluating not only daily precipitation amount, for which the method was originally designed (Themeßl et al. 2011), but also daily mean, minimum, and maximum temperature.

We apply a decadal "leave one out" cross validation approach (e.g., Themeßl et al. 2011), where each decade within 1961–2000 is corrected independently with the remaining 30 years used for calibration. The error characteristics are discussed either averaged over the four validation decades via spatial bias maps and tables or for the entire 40 validation years via *pdfs* for sub-regions.

Figure 4 exemplarily shows the seasonal bias pattern of the uncorrected and corrected CCLM precipitation amount in winter (DJF) and summer (JJA) over Europe. Table 1 complements Fig. 4 summarizing the respective sub-regional mean biases for all seasons as well as the observed values. With the exception of JJA, CCLM is too wet with biases between −0.7 mm/d to +1.5 mm/d (about −40% and +70%). In DJF the bias pattern features a pronounced spatial variation and particularly follows orographic structures. The same pattern can be found in spring (MAM) and autumn (SON) (not shown). Referring to Hofstra et al. (2009), who concluded that E-OBS is of limited quality in areas with complex terrain, these biases could be partly due to errors in the reference dataset. In summer, precipitation is strongly underestimated in southern Europe with dry maxima in IP, MD and EA.

With QM (including frequency adaptation) the bias is almost removed across Europe, independently of region and season. Remaining absolute biases amount to ≤0.1 mm/d at sub-regional scale and ≤0.5 mm/d at grid cell scale.

The *pdfs* of daily precipitation amount in Fig. 5 are shown for sub-regions IP, SC, and AL. These sub-regions have been selected for closer evaluation due to their differing climatic conditions and their differing bias characteristics and will be used subsequently for all further analyzes. The *pdfs* reveal a distinct intensity dependency of

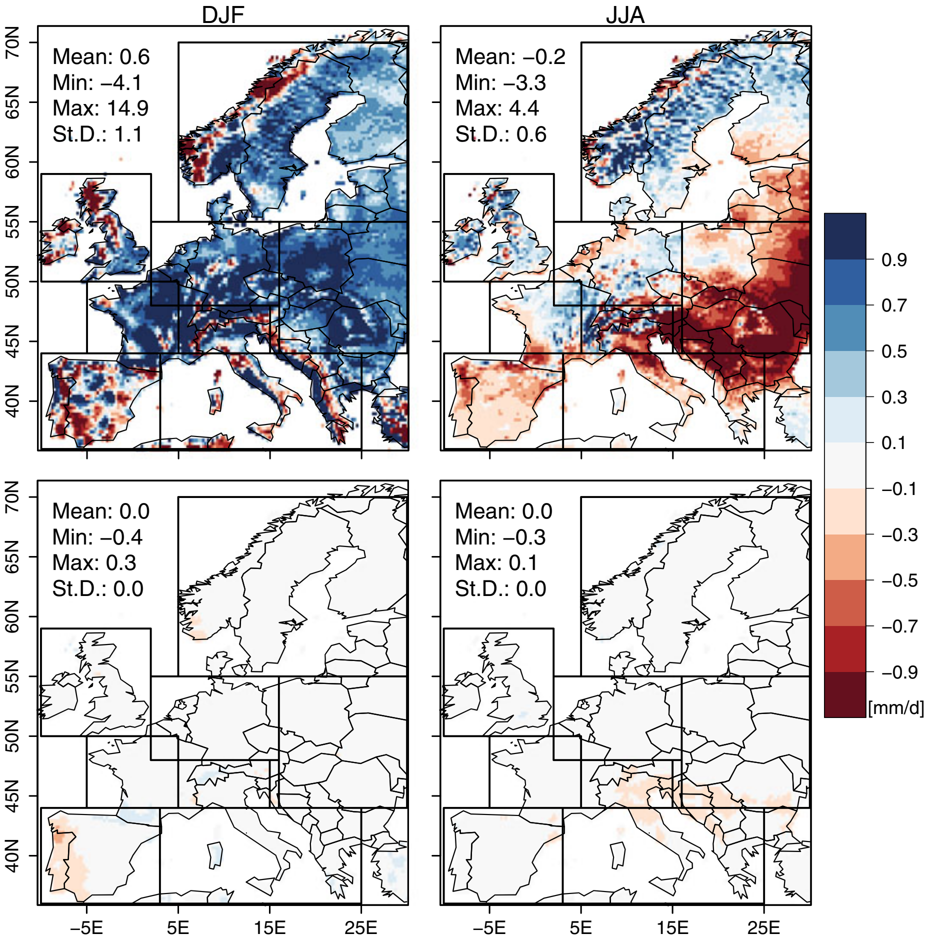

Fig. 4 Summer (JJA) and winter (DJF) seasonal bias of uncorrected (*upper two panels*) and corrected (*lower two panels*) precipitation amount from the RCM control simulation (1961–2000) compared to E-OBS. Statistics for entire Europe are given in the upper left corner of each panel

the precipitation error. While the dry-day frequency is systematically underestimated by CCLM, the frequency of light precipitation events between 0.1 mm/d and 1 mm/d ("drizzling") as well as of heavy precipitation events are mostly overestimated by CCLM.

Quantile mapping corrects the frequency of dry days adequately, also through the implementation of FA, and properly adjusts intensities below 30 mm/d; the frequency of higher intensities (usually above the 99 percentile) is still often overestimated, depending on season and region. However, these remaining errors are in general smaller than the RCM errors.

The temperature evaluation (Fig. 6 and Table 1) reveals that the RCM is cold biased across Europe in all seasons but JJA. The respective biases range from −3.6 K to +2.8 K on sub-regional scale. Smaller scale biases are larger and often associated with orographical features and coastlines. In summer the model exhibits a strong warm bias in large parts of

Table 1 Sub-regional and seasonal biases of the uncorrected RCM (RCM), the corrected RCM (QM), and the respective observational data (E-OBS) for mean temperature [degC] and precipitation amount [mm/d]

Sub-region	DJF			MAM			JJA			SON		
	E-OBS	RCM	QM	E-OBS	RCM	QM	E-OBS	RCM	QM	E-OBS	RCM	QM
	Precipitation amount											
BI	3.4	0.2	0.0	2.4	0.2	0.0	2.4	0.2	0.0	3.4	−0.1	0.0
IP	2.6	0.2	−0.1	1.9	0.1	0.1	0.8	−0.3	0.0	2.2	0.0	0.0
FR	2.3	1.1	0.0	2.0	0.7	0.0	1.7	0.1	0.0	2.4	0.5	0.0
ME	1.8	0.9	0.0	1.8	0.6	0.0	2.2	0.0	0.0	1.9	0.4	0.0
SC	1.7	0.5	0.0	1.3	0.4	0.0	2.2	0.2	0.0	2.3	0.3	0.0
AL	2.6	1.5	0.0	3.0	1.0	0.1	3.2	−0.3	−0.1	3.3	0.9	0.1
MD	2.2	0.7	0.0	1.7	0.7	0.0	1.0	−0.4	−0.1	2.2	1.0	0.1
EA	1.1	0.8	0.0	1.4	0.7	0.0	2.2	−0.7	0.0	1.5	0.6	0.0
	Mean temperature											
BI	3.9	−0.5	0.0	7.6	−1.2	−0.1	14.1	−0.4	0.0	9.4	−0.3	0.0
IP	6.5	−0.7	0.0	11.6	−0.9	−0.1	20.8	1.7	0.0	14.1	0.0	0.0
FR	4.4	−0.4	0.0	9.9	−1.5	−0.1	17.8	1.0	0.0	11.7	−0.4	0.0
ME	1.1	−0.8	0.0	8.3	−1.7	−0.1	16.8	0.9	0.0	9.3	−0.9	0.0
SC	−7.7	−1.2	−0.1	1.3	−3.6	−0.2	13.0	−1.1	0.0	2.8	−1.0	0.1
AL	−0.2	−1.6	0.0	6.9	−2.6	−0.1	16.1	1.0	0.0	8.6	−1.1	0.0
MD	4.7	−0.9	−0.1	11.1	−1.2	−0.1	20.9	2.3	0.0	13.7	−0.2	0.0
EA	−2.0	−1.0	−0.1	8.3	−1.9	−0.1	18.0	2.8	0.0	8.7	−0.6	0.0

Europe peaking in IP, EA, and MD. In combination with the dry bias, this deviation represents the already mentioned summer drying problem, which is found for several RCMs over Europe. In SC, the RCM is cold biased.

Biases of minimum and maximum temperature partly strongly deviate in their patterns as well as magnitude from mean temperature biases (not shown).

Regardless of the spatially and temporally varying error characteristics, QM corrects the temperature bias to virtually zero throughout Europe. Remaining absolute biases are smaller or equal 0.2 K on the sub-regional scale and smaller or equal 0.3 K on the grid cell scale. Similar results are obtained for daily minimum and maximum temperature (not shown).

The uncorrected RCM *pdfs* in Fig. 7 roughly capture the seasonal characteristics but show notable deviations from the observational reference. The errors at different percentiles are mostly in the range of ±3 K and highly magnitude dependent. Remarkable is the frequency peak of the RCM at zero degree. This is most probably related to problems in the representation of melting and freezing processes in the CCLM soil and snow models.

In contrast, all corrected *pdfs* nicely resemble the observations and do not feature any spurious deviations at zero degree. However, the percentile difference plots also show that the tails of the corrected distributions still feature considerable errors. Nevertheless, they tend to be smaller than the uncorrected ones and only concern rare values outside the $\pm 2\sigma$ range. Similar results are obtained for daily minimum and daily maximum temperature (not shown).

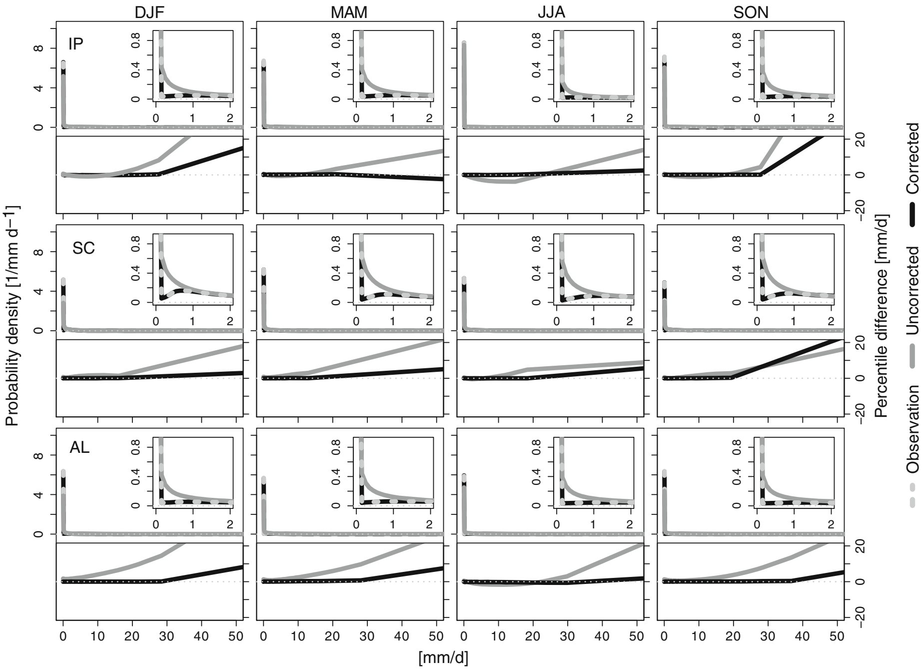

Fig. 5 Seasonal observed, uncorrected model, and corrected model *pdfs* (*upper sub-panels*) of precipitation amount for IP (*top row*), SC (*middle row*) and AL (*bottom row*). Differences between the uncorrected/corrected and observed data at different percentiles are shown in the *lower sub-panels*

3.2 Evaluation of QM with regard to derived parameters

The performance of QM with regard to derived parameters is investigated in this subsection. For this purpose, we consider different indices, which are listed and defined in Table 2.

The RCM generally overestimates precipitation-related indices (Figure S1 in the Electronic Supplementary Material and Table 3) over most parts of Europe. The error patterns for pint and pn10 essentially show the same structure with large areas ≥+0.7 mm/d and ≥+7 days/year (days/y), respectively and orographically induced underestimation. Px1d is overestimated even more widespread by more than +9 mm/d.

QM reduces the sub-regional absolute biases to ≤0.1 mm/d (≤0.5 mm/d on grid cell scale) for pint, ≤0.3 days/y (≤2.6 days/y on grid cell scale) for pn10, and ≤4.4 mm/d (≤19.5 mm/d on grid cell scale) for px1d. QM furthermore adapts the shape and kurtosis of the *pdfs* to a very large degree (exemplarily shown for AL in Fig. 8 but similar results are obtained for IP and SC). However, all three corrected indices of precipitation extremes feature a slight wet bias at the lowest values (similar to the uncorrected indices). Furthermore, a reduced overestimation of px1d remains above about 75 mm/d.

All uncorrected temperature extreme indices in Figure S2 in the Electronic Supplementary Material and Table 3 are warm biased in average and feature a north−south gradient. For tasx the sub-regional biases vary between −0.6 K and +5.2 K, whereas sub-regional biases for txn25 and tnn20 amount up to about +28 days/y and +20 days/y, respectively. In the case of tnn20, the maximum biases exceed the respective observations by a factor of more than two.

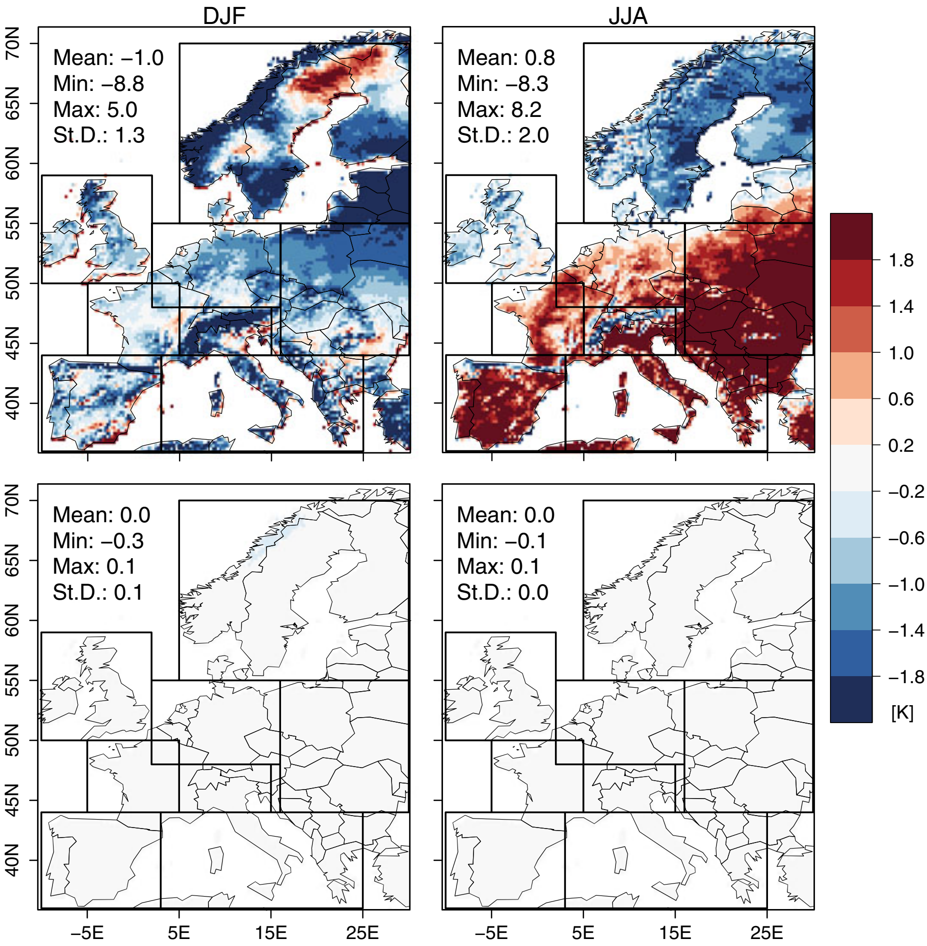

Fig. 6 The same as in Fig. 4 but for mean temperature

After error-correction the biases are reduced by roughly one order of magnitude: Remaining absolute biases are ≤0.6 K (≤1.6 K on the grid cell scale) for tasx, ≤2.2 days/y (≤4.4 days/y on grid cell scale) for txn25, and ≤0.4 days/y (≤2.7 days/y on grid cell scale) for tnn20. The errors in the shape as well as in the kurtosis of the *pdfs* of the temperature extremes are strongly reduced (see Fig. 8). Minor discrepancies at the tail of the tasx distribution (smaller than ±3 K) remain, whereas errors for corrected txn25 and tnn20 remain small throughout the entire distribution. Similar results are obtained for IP and SC (not shown).

3.3 Impact of QM on the climate change signal

In order to evaluate the impact of QM on the CCS, QM calibrated on the control simulation between 1961 and 2000 is applied to the RCM control and scenario simulation until 2050. CCSs are calculated on monthly basis between the periods 1971–2000 and 2021–2050 for the uncorrected as well as the corrected time series and compared with regard to the mean

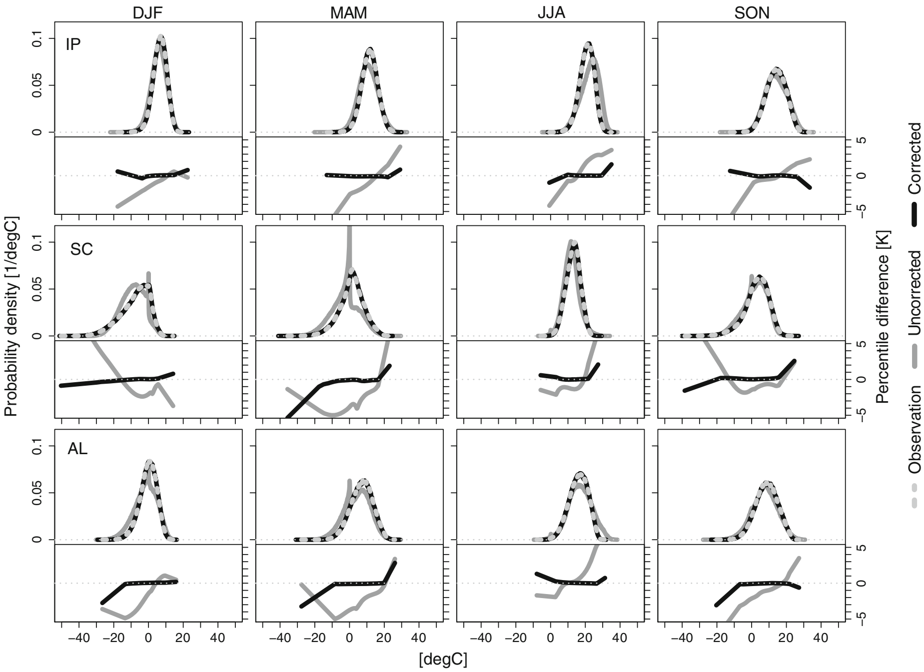

Fig. 7 The same as in Fig. 5 but for mean temperature

CCS, spatial CCS pattern, annual cycle of the CCS, and the significance of the monthly CCS. In cases with small CCSs numbers are given with higher accuracy. The significance of the climate change signals is determined by the Wilcoxon rank sum test (Wilks 1995) on the 95% significance level.

Figure 9 displays uncorrected and corrected CCSs of precipitation amount and mean temperature over Europe. Table 4 lists the respective sub-regional mean numbers. The uncorrected annual precipitation CCS shows a north-south gradient over Europe from an increase of +0.3 mm/d in SC to a decrease of −0.2 mm/d in IP. This pattern is already known from various RCM simulations and, e.g., discussed in Giorgi and Coppola (2007) or van der Linden and Mitchell (2009). QM leads to scattered local impact on the CCS mostly around mountain ridges and coastlines and only small impact on the sub-regional scale. IP, SC, and AL feature distinct annual cycles of the CCSs, which vary between −0.7 mm/d and

Table 2 List of analyzed derived indices of extremes with their abbreviation, long name, and definition

Acronym	Long name	Definition
tasx	Maximum mean temperature	Maximum of daily mean temperature
txn25	No. of summer days	No. of days with daily maximum temperature >25°C
tnn20	No. of tropical nights	No. of days with daily minimum temperature >20°C
pint	Precipitation intensity	Mean precipitation amount on days ≥1 mm/d
pn10	Heavy precipitation days	No. of days with precipitation amount ≥10 mm/d
px1d	Maximum 1 day precipitation	Maximum of daily precipitation amount

Table 3 As in Table 1 but on annual basis for pint [mm/d], pn10 [days/y], px1d [mm/d], tasx [K], txn25 [days/y], and tnn20 [days/y]

Sub-region	E-OBS	RCM	QM	E-OBS	RCM	QM	E-OBS	RCM	QM
	pint			pn10			px1d		
BI	5.9	0.1	0.0	28.9	0.5	−0.1	32.9	4.0	2.0
IP	7.2	0.2	0.1	20.8	−2.4	0.0	35.1	13.2	3.4
FR	5.5	0.7	0.0	18.6	6.7	0.0	28.1	14.4	2.4
ME	4.7	0.6	0.0	13.9	5.9	0.0	25.2	11.9	2.9
SC	4.6	0.6	0.0	13.4	4.7	0.0	23.5	8.9	1.7
AL	8.5	0.6	0.1	35.6	5.2	−0.3	47.1	21.8	4.4
MD	6.9	1.5	0.1	19.4	2.8	−0.1	33.7	27.8	4.2
EA	4.7	0.7	0.1	11.4	4.4	−0.1	25.5	11.2	2.8
	tasx			txn25			tnn20		
BI	19.4	0.2	0.0	3.8	1.3	0.0	0.0	0.1	0.0
IP	27.2	2.2	−0.3	89.6	24.1	−2.1	4.0	15.3	−0.1
FR	24.7	2.5	−0.5	37.2	15.6	−0.9	0.3	5.4	0.0
ME	24.2	2.3	−0.5	27.0	15.1	−0.6	0.1	2.1	0.0
SC	20.0	−0.6	−0.3	6.6	0.7	0.0	0.0	0.1	0.0
AL	22.2	2.0	−0.6	32.9	15.3	−0.9	1.9	6.6	−0.1
MD	26.9	2.7	−0.1	82.5	24.4	−2.2	8.8	20.1	−0.4
EA	24.5	5.2	−0.3	44.9	27.7	−1.4	0.4	12.3	0.0

+0.6 mm/d (Fig. 9, upper rightmost panel). QM does not change the general characteristics of these annual cycles and the respective differences remain below 0.1 mm/d. E.g., the larger negative CCS in AL in summer is related to more wet days ≤50 mm/d in the control period compared to the scenario period and the higher positive correction values at these intensities (Figure S3 in the Electronic Supplementary Material). In almost all cases the significance of the obtained CCSs remains unchanged by QM.

Regarding temperature, the uncorrected simulation yields an increase of +2.0 K on average over Europe until the mid of the 21st century, with most sub-regional differences of about ±0.5 K (compare Table 4). Similar to precipitation amount, QM only moderately changes the annual CCS, with sub-regional absolute impact smaller or equal 0.3 K. The impact pattern features a north-south gradient from local increase in northern SC to more widespread decreases in EA, AL, MD, and IP. The corresponding annual cycles in IP, SC, and AL accord with these moderate changes of QM. Those impacts in single months, such as in January in AL can be explained with the help of Figure S3 in the Electronic Supplementary Material and similar reasoning as for precipitation amount.

Figure S4 in the Electronic Supplementary Material shows the respective results for pint, pn10 and px1d. Precipitation intensity (pint) shows a rather homogenous increase in the uncorrected simulation with large areas ≥+0.2 mm/d. In contrast, px1d and especially pn10 exhibit similar north-south gradients as obtained for precipitation amount. In both cases, the positive signals in SC and BI exceed the other sub-regional European CCSs by about 100% in absolute terms (compare Table 4). The impact of QM on the annual CCS of all precipitation related indices is scattered and exceeds 50% in rare cases on the sub-regional scale. QM modifies the annual cycles stronger than the respective mean precipitation

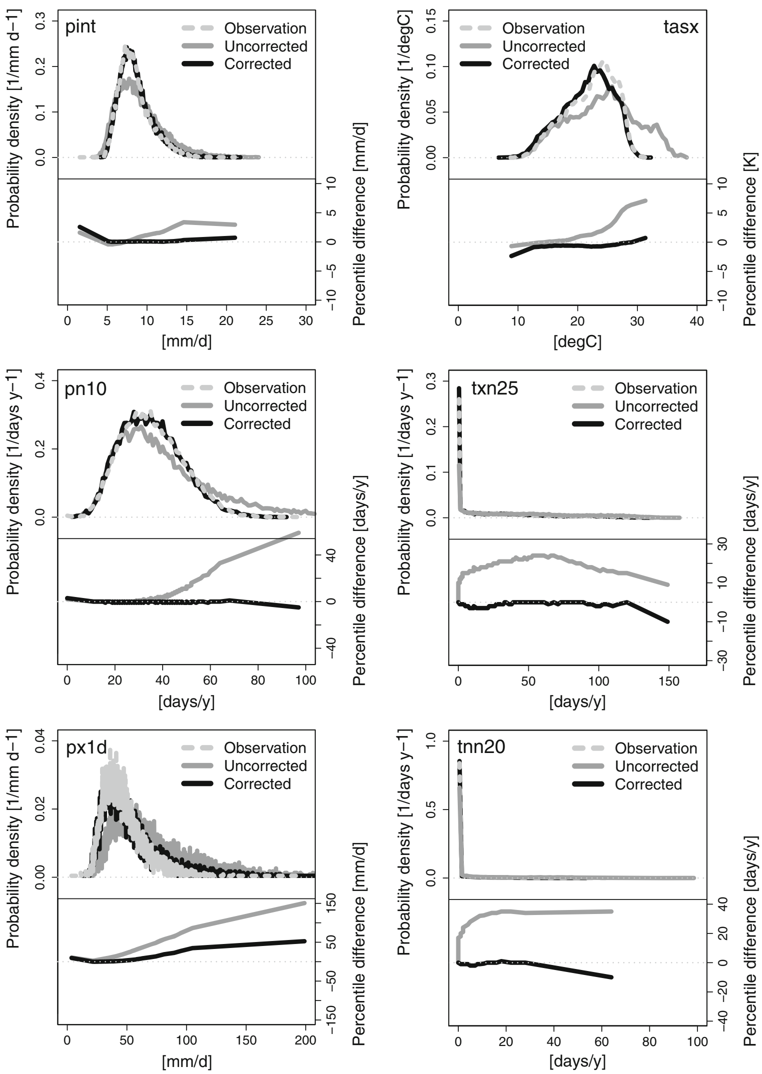

Fig. 8 Annual *pdfs* of the uncorrected model, the corrected model, and observations for pint, pn10, px1d, tasx, txn25, and tnn20 in AL. The *lower sub-panels* present the corresponding differences between the uncorrected/corrected and observed data at different percentiles

cycles, although the significances remain unchanged. For example, the CCS of pint in IP in July is reduced by about 60%, which consequently alters the annual cycle by shifting the respective maximum to September. Furthermore, in the case of pint and px1d in IP, QM increases the autumn CCS (up to 200% of the uncorrected CCS), which converts single negative months to positive ones.

Figure 10, finally, presents uncorrected and corrected CCSs of temperature-related extremes. Uncorrected tasx shows a similar CCS pattern as mean temperature but a smaller areal average increase of +1.8 K. Uncorrected txn25 and tnn20, on the other hand, feature a gradient of increasing summer days and tropical nights from north to south, with no change northern of a certain latitude. The southern boundary of this zero-change area is strongly

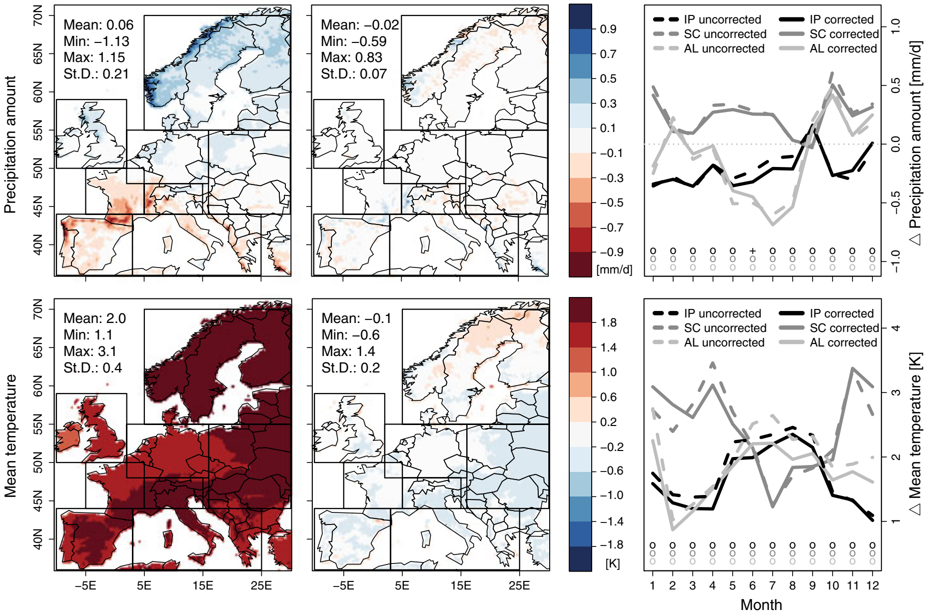

Fig. 9 Annual mean maps of the uncorrected monthly CCS (*left column*), the difference between the uncorrected and the corrected CCS (*middle column*), and the respective annual cycles of the CCS for three sub-regions. *Top row*: precipitation amount; *bottom row*: mean temperature. In the lower part of the annual cycle plots change of significance is indicated with “o” (unchanged significance), “–” (loss of significance after correction), and “+” (significance established after correction)

related to the threshold values used in the calculation of the indices. The sub-regional maximum changes amount to +1.8 days/month (days/m) (≙21.6 days/y) and +1.7 days/m (≙20.4 days/y) for txn25 and tnn20, respectively (compare Table 4).

The impact of QM on the annual CCS is not uniform between indices and regions. The sub-regional impact of QM on tasx CCS is around 0.2 K or smaller, whereas QM modifies the CCS of txn25 and tnn20 by up to 80%.

While the uncorrected sub-regional annual cycles of the CCSs of tasx as well as the impacts of QM are comparable to mean temperature, the annual cycles of txn25 and tnn20 yield different results. In the case of txn25 strong impacts are visible for SC and IP, while the scattered local impact cancels out in AL. In IP in August, the CCS is more than doubled by QM and the maximum is shifted from spring to autumn. This drastic reduction can mainly be related to the correction of a positive bias in maximum temperature by QM (not shown). In combination with a positive trend in maximum temperature, this bias reduction results in far stronger reduction of threshold exceedances in the control than in the scenario period. In contrast, QM drastically reduces the monthly CCSs of tnn20 for IP and AL up to about 60% due to the same reasons as for txn25. Concerning the changes in the significance of the CCSs for temperature extremes, no systematic impacts are obtained.

4 Summary and conclusions

This study evaluates the performance of an empirical-statistical downscaling and error correction method (DECM), quantile mapping (QM), applied to a RCM climate simulation

Table 4 Sub-regional annual CCS of precipitation amount (prec) [mm/d], mean temperature (temp) [K], as well as of the derived extremes (pint [mm/d], pn10 [days/m], px1d [mm/d], tasx [K], txn25 [days/m], and tnn20 [days/m]). RCM indicates the uncorrected CCS and QM the difference between the uncorrected and the corrected CCS

Sub-region	RCM	QM	RCM	QM	RCM	QM	RCM	QM
	prec		temp		pint		pn10	
BI	0.11	0.00	1.5	−0.1	0.35	−0.01	0.24	0.00
IP	−0.20	−0.02	1.8	−0.1	0.24	−0.06	−0.16	−0.04
FR	−0.18	0.02	1.7	−0.1	0.09	0.00	−0.11	0.01
ME	0.04	0.00	1.7	−0.1	0.22	−0.03	0.12	−0.03
SC	0.26	−0.03	2.4	0.0	0.39	−0.07	0.32	−0.08
AL	−0.11	0.00	2.0	−0.2	0.15	−0.02	−0.12	−0.01
MD	−0.10	−0.01	1.8	−0.1	0.28	−0.11	−0.11	0.01
EA	0.02	−0.05	1.9	−0.3	0.19	−0.07	0.04	−0.06
	px1d		tasx		txn25		tnn20	
BI	1.21	−0.03	1.2	0.1	0.21	0.03	0.01	−0.01
IP	−0.62	0.04	1.9	0.0	1.62	0.39	1.64	−0.71
FR	−0.37	0.20	1.8	0.0	1.80	0.04	0.78	−0.57
ME	0.62	0.04	1.6	0.0	0.95	−0.09	0.35	−0.28
SC	1.61	−0.32	2.0	0.0	0.15	0.13	0.02	−0.01
AL	0.04	0.06	2.1	−0.1	1.44	−0.07	0.76	−0.32
MD	0.09	−0.13	1.7	0.2	1.50	0.45	1.74	−0.40
EA	0.38	−0.09	1.8	−0.2	0.99	0.12	0.81	−0.60

over Europe regarding daily mean, minimum and maximum temperature, daily precipitation amount, and derived indices of extremes. In addition, two issues related to the climate change context are discussed in more detail: A methodological extension of QM which allows "new extremes" (values outside the calibration range) and the impact of QM on the climate change signal (CCS).

In a decadal cross validation of a 40 years RCM (CCLM) control simulation, QM confirms its applicability to longer climate simulations and to several parameters, regardless of spatially and temporally varying error characteristics. This regional transferability strongly suggests a general transferability of QM to any regional climate model, which is also underpinned by results from Déqué (2007; applied to a variable resolution GCM), Piani et al. (2009), Themeßl et al. (2011), Bardossy and Pegram (2011), Heinrich and Gobiet (2011), or Rojas et al. (2011).

Quantitatively, we demonstrate that QM reduces biases of daily mean, minimum, maximum temperature, and daily precipitation amount by roughly one order of magnitude. In most cases, the remaining absolute biases are smaller or equal 0.1 K for temperature and smaller or equal 0.1 mm/d for precipitation amount on the sub-regional scale. For daily precipitation these results are obtained only after frequency adaptation, which assures an adequate performance of QM in situations with more modeled than observed dry days.

Concerning derived indices of extremes, QM shows comparable skill as for daily temperature and precipitation. Particularly indices related to threshold values can feature tremendous biases in uncorrected RCM data, which can be easily removed by

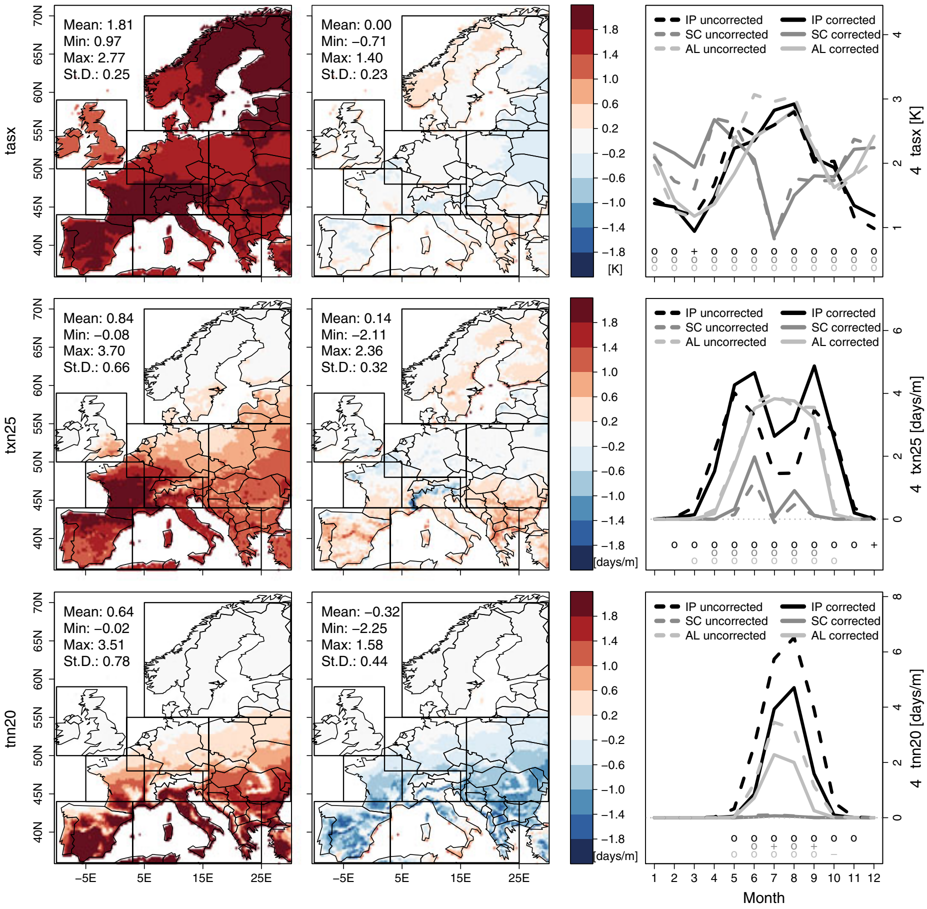

Fig. 10 Same as Fig. 9 but for tasx (*top row*), txn25 (*middle row*) and tnn20 (*bottom row*)

QM. Nevertheless, absolute extremes are more prone to biases after error correction than mean parameters. The analyzed indices do not include parameters which are based on temporal persistence characteristics such as maximum number of consecutive dry days.

Errors in the shapes of the daily temperature and precipitation *pdfs* are corrected adequately at least within the ±2σ range around the mean (temperature) or up to the 99th quantile (precipitation). Regarding "new extremes", it could be demonstrated that by simple extrapolation of the correction terms QM successfully produces new extremes without deterioration (and mostly with improvement) of the RCM quality.

Applied to future scenarios, this study shows that QM can moderately modify the climate change signal (CCS) of the corrected parameters. These impacts can be explained by the combination of two factors: a magnitude-dependent error correction functions, thus low and high quantities are corrected differently, and a trend in the underlying data, thus uncorrected future periods feature a significantly changed *pdf* compared to the respective baseline. Similar findings are reported by Hagemann et al. (2011). Furthermore, CCSs of indices that are non-linearly derived from the corrected quantities, such as threshold indices, can be strongly modified by QM. In particular, the altered CCSs of derived threshold indices, such as the number of summer days or tropical nights, are regarded as more reliable than the respective uncorrected CCSs. This is due to the fact that in these cases the impact of QM is mainly related to a partly drastic reduction of errors in the position of the *pdf*. Without correction, this error leads to a non-linear overestimation of such indices.

However, in general the reliability of corrected scenarios strongly depends on the source of the error in the RCMs and on the stationarity of the error correction functions. Concerning the former, QM will likely increase the reliability in the constructed scenarios if the error is related to the shape of the distribution, thus magnitude-dependent. In contrast, errors in temporal characteristics stemming from an erroneous sequence of circulation systems in the climate model are yet not improved by QM as e.g., a warm day still remains warm after correction. However, QM does not degrade such temporal characteristics (compare Déqué 2007; Themeßl et al. 2011). Furthermore, a study by Srikanthan and Pegram (2009) shows that improvement of QM with regard to temporal characteristics seems feasible.

Concerning the issue of stationarity, it has to be noted that for the application of QM to single decades errors are expected to be larger due to the yet not fully investigated impact of decadal climate variability on the stationarity of DECMs. Studies such as by Salathé (2005), who shows that a GCM correction calibrated on Pacific Decadal Oscillation (PDO) warm phase performs equally well on cool PDO epochs, however indicate a general stability in the correction process. In addition, it can be expected that the stationarity restriction affects our application to a lesser extent than most other statistical downscaling approaches, since in our application no considerable scale gaps have to be bridged and since atmospheric parameters are directly mapped. By this means, the effects of climate change are already represented in the predictor to a large degree and the transfer function can be expected to be largely unaltered by climate change within the calibration range. Given that, and the demonstrated skill of QM within the calibration range and partly beyond, we are confident in the application of QM to future scenarios. However, a consistent quantification of the limitations arising from the stationary assumption remains an important issue for further research in our field.

In application to climate change impact investigations it should be kept in mind that QM, as applied here, post-processes each variable separately. As a consequence, the physical consistence between variables and/or autocorrelation structure may be distorted, which could lead to unexpected effects in the impact models (e.g., Boé et al. 2007).

Acknowledgements The ENSEMBLES data used in this work was funded by the EU FP6 Integrated Project ENSEMBLES (Contract number 505539) whose support is gratefully acknowledged especially the CCLM simulations provided by the ETH Zurich. We furthermore acknowledge the E-OBS dataset from the ENSEMBLES project and the data providers in the ECA&D project (http://eca.knmi.nl) as well as the EU FP6 project CLAVIER which partly funded this study.

References

Bardossy A, Pegram G (2011) Downscaling precipitation using regional climate models and circulation patterns toward hydrology. Water Resour Res 47:W04505. doi:10.1029/2010WR009689

Benestad RE, Hanssen-Bauer I, Chen D (2008) Empirical statistical downscaling. World Scientific Publishing Company, New Jersey

Boé J, Terray L, Habets F, Martin E (2007) Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies. Int J Climatol 27:1643–1655
Böhm U, Kücken M, Ahrens W, Block A, Hauffe D, Keuler, K Rockel B, Will A (2006) Clm—the climate version of lm: brief description and long-term applications. COSMO Newsletter, 6
Christensen JH, Christensen OB (2007) A summary of the PRUDENCE model projections of changes in European climate by the end of this century. Clim Chang 81:7–30. doi:10.1007/s10584-006-9210-7
Déqué M (2007) Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: model results and statistical correction according to observed values. Glob Planet Chang 57:16–26
Dettinger MD, Cayan DR, Meyer MK, Jeton AE (2004) Simulated hydrologic responses to climate variations and change in the Merced, Carson, and American river basins, Sierra Nevada, California, 1900–2099. Clim Chang 62:283–317
Dobler A, Ahrens B (2008) Precipitation by a regional climate model and bias correction in Europe and South Asia. Meteorol Z 17:499–509
Fowler HJ, Kilsby CG (2007) Using regional climate model data to simulate historical and future river flows in northwest England. Clim Chang 80:337–367
Frei C, Christensen JH, Déqué M, Jacob D, Jones RG, Vidale PL (2003) Daily precipitation statistics in regional climate models: evaluation and intercomparison for the European Alps. J Geophys Res 108 (D3):4124. doi:10.1029/2002JD002287
Giorgi F, Coppola E (2007) European climate-change oscillation (ECO). Geophys Res Lett 34:L21703. doi:10.1029/2007GL031223
Giorgi F, Mearns LO (1991) Approaches to the simulation of regional climate change: a review. Rev Geophys 29:191–216
Giorgi F, Mearns LO (1999) Introduction to special section: regional climate modelling revisited. J Geophys Res 104:6335–6352
Gordon C, Cooper C, Senior CA, Banks H, Gregory JM, Johns TC, Mitchell JFB, Wood RA (2000) The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. Clim Dyn 16:147–168
Graham LP, Andréasson J, Carlsson B (2007) Assessing climate change impacts on hydrology from an ensemble of regional climate models, model scales and linking methods—a case study on the Lule River basin. Clim Chang 81:293–307
Gutowski WJ, Decker SG, Donavon RA, Pan Z, Arritt RW, Takle ES (2003) Temporal-spatial scales of observed and simulated precipitation in central U.S. climate. J Clim 16:3841–3847
Hagemann S, Machenhauer B, Jones R, Christensen OB, Déqué M, Jacob D, Vidale PL (2004) Evaluation of water and energy budgets in regional climate models applied over Europe. Clim Dyn 23:547–567
Hagemann S, Chen C, Haerter JO, Heinke J, Gerten D, Piani C (2011) Impact of a statistical bias correction on the projected hydrological changes obtained from three GCMs and two hydrology models. J Hydrometeorol, in press
Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New M (2008) A European daily high-resolution gridded dataset of surface temperature and precipitation. J Geophys Res 113:1–12, D20119. doi:10.1029/2008JD10201
Heinrich G, Gobiet A (2011) The future of dry and wet spells in Europe: a comprehensive study based on the ENSEMBLES regional climate models. Int J Climatol, (in press)
Hofstra N, Haylock M, New M, Jones PD (2009) Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature. J Geophys Res 114:D21101. doi:10.1029/2009JD011799
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 inter-comparison of regional climate models for Europe: design of the experiments and model performance. Clim Chang 81:31–52. doi:10.1007/s10584-006-9213-4
Leander R, Buishand TA (2007) Resampling of regional climate model output for the simulation of extreme river flows. J Hydrol 332:487–496
Leander R, Buishand TA, van den Hurk BJJM, de Wit MJM (2008) Estimated changes in flood quantiles of the river Meuse from resampling of regional climate models. J Hydrol 351:331–343
Lenderik G, Buishand A, van Deursen W (2007) Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach. Hydrol Earth Syst Sci 11:1145–1159
Maraun D, Ireson A, Wetterhall F, Bachner S, Kendon E, Rust HW, Venema VKC, Widmann M, Chandler RE, Onof CJ, Osborn TJ, Sautner T, Themeßl M, Thiele-Eich I (2010) Statistical downscaling and modelling of precipitation. Bridging the gab between dynamical models and the end users. Rev Geophys 48:RG3003. doi:10.1029/2009RG000314
Nakicenovic 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, Raihi K, Roehrl A,

Rogner H-H, Sankovski A, Schlesinger M, Shukla P, Smith S, Swart R, van Rooijen S, Victor N, Dadi Z (2000) IPCC Special Report on Emissions Scenarios. Cambridge University Press, Cambridge

Piani C, Haerter JO, Coppola E (2009) Statistical bias correction for daily precipitation in regional climate models over Europe. Theor Appl Climatol. doi:10.1007/s00704-009-0134-9

Piani C, Weedon GP, Best M, Gomes SM, Viterbo P, Hagemann S, Haerter JO (2010) Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. J Hydrol. doi:10.1016/j.jhydrol.2010.10.024

Rojas R, Feyen L, Dosio A, Bavera D (2011) Improving pan-european hydrological simulation of extreme events through statistical bias correction of RCM-driven climate simulations. Hydrol Earth Syst Sci Discuss 8:3883–3936

Salathé EP (2005) Downscaling simulations of future global climate with application to hydrologic modeling. Int J Climatol 25:419–436

Srikanthan R, Pegram G (2009) A nested multisite daily rainfall stochastic model generator. J Hydrol 371:142–153

Suklitsch M, Gobiet A, Leuprecht A, Frei C (2008) High resolution sensitivity studies with the regional climate model CCLM in the Alpine Region. Meteorol Z 17:467–476

Suklitsch M, Gobiet A, Truhetz H, Awan NK, Göttel H, Jacob D (2010) Error characteristics of high resolution regional climate models over the Alpine area. Clim Dyn. doi:10.1007/s00382-010-0848-5

Themeßl M, Gobiet A, Leuprecht A (2011) Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. Int J Climatol 31:1531–1544. doi:10.1002/joc.2168

Uppala SM, Kållberg PW, Simmons AJ, Andrae U, Da Costa BV, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, van De Berg L, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hoskins BJ, Isaksen L, Janssen PAEM, Jenne R, Mcnally AP, Mahfouf JF, Morcrette JJ, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J (2005) The ERA-40 re-analysis. Q J R Meteorol Soc 131(612):2961–3012

van der Linden P, Mitchell JFB (2009) ENSEMBLES: climate change and its impacts: summary of research and results from the ENSEMBLES project. Met Office Hadley Centre, Exeter

van Pelt SC, Kabat P, ter Maat HW, van den Hurk BJJM, Weerts AH (2009) Discharge simulation performed with a hydrological model using bias corrected regional climate model input. Hydrol Earth Syst Sci 13:2387–2397

Wang Y, Leung LR, McGregor JL, Lee DK, Wang WC, Ding Y, Kimura F (2004) Regional climate modeling: progress, challenges, and prospects. J Meteorol Soc Japan 82:1599–1628

Wilks DS (1995) Statistical methods in atmospheric science, volume 59 of International Geophysics Series. Academic, San Diego

Wood AW, Leung LR, Sridhar V, Lettenmaier DP (2004) Hydrologic implications of dynamical and statistical approaches to downscale climate model outputs. Clim Chang 62:189–216