

# Extended regional climate model projections for Europe until the mid-twentyfirst century: combining ENSEMBLES and CMIP3

Georg Heinrich · Andreas Gobiet · Thomas Mendlik

Received: 30 August 2012 / Accepted: 10 June 2013 / Published online: 25 June 2013  
© The Author(s) 2013. This article is published with open access at Springerlink.com

**Abstract** This study aims at sharpening the existing knowledge of expected seasonal mean climate change and its uncertainty over Europe for the two key climate variables air temperature and precipitation amount until the mid-twentyfirst century. For this purpose, we assess and compensate the global climate model (GCM) sampling bias of the ENSEMBLES regional climate model (RCM) projections by combining them with the full set of the CMIP3 GCM ensemble. We first apply a cross-validation in order to assess the skill of different statistical data reconstruction methods in reproducing ensemble mean and standard deviation. We then select the most appropriate reconstruction method in order to fill the missing values of the ENSEMBLES simulation matrix and further extend the matrix by all available CMIP3 GCM simulations forced by the A1B emission scenario. Cross-validation identifies a randomized scaling approach as superior in reconstructing the ensemble spread. Errors in ensemble mean and standard deviation are mostly less than 0.1 K and 1.0 % for air temperature and precipitation amount, respectively. Reconstruction of the missing values reveals that expected seasonal mean climate change of the ENSEMBLES RCM projections is not significantly biased and that the associated uncertainty is not underestimated due to sampling of only a few driving GCMs. In contrast, the spread of the

extended simulation matrix is partly significantly lower, sharpening our knowledge about future climate change over Europe by reducing uncertainty in some regions. Furthermore, this study gives substantial weight to recent climate change impact studies based on the ENSEMBLES projections, since it confirms the robustness of the climate forcing of these studies concerning GCM sampling.

**Keywords** Future climate projections · Regional climate models · ENSEMBLES · Uncertainty · Sampling bias · Europe

## 1 Introduction

The application of general circulation models (GCMs) driven by prescribed greenhouse gas (GHG) emission scenarios is nowadays the most common way to obtain physically based climate projections. Due to their coarse spatial resolution (typically 100–300 km horizontal grid spacing with an effective resolution of about 1,000 km) GCMs currently fail to properly represent many regional and local climate processes, such as, e.g., orographic precipitation (McGregor 1997). In order to generate climate simulations with finer horizontal resolution, regional climate models (RCMs) are nested within the GCM large scale atmospheric circulation over a limited area (Giorgi and Mearns 1991, 1999; McGregor 1997; Wang et al. 2004; Rummukainen 2010) and the added value of RCMs in representing regional climate characteristics has been demonstrated in several studies (Jones et al. 1995; Laprise 2003; Castro et al. 2005; Buonomo et al. 2007; Feser et al. 2011).

Regional as well as global climate projections are subject to considerable uncertainties which can be roughly

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s00382-013-1840-7) contains supplementary material, which is available to authorized users.

---

G. Heinrich (✉) · A. Gobiet · T. Mendlik  
Wegener Center for Climate and Global Change (WEGC)  
and Institute for Geophysics, Astrophysics,  
and Meteorology/Institute of Physics (IGAM/IP),  
University of Graz, Graz, Austria  
e-mail: g.heinrich@uni-graz.at

divided into three components: (1) Uncertainty due to natural variability, (2) uncertainty due to unknown future GHG emissions, and (3) uncertainty due to imperfect simulation of the climate system (Collins 2007). In order to analyse these uncertainties in RCM projections, large-scale European projects such as PRUDENCE (<http://prudence.dmi.dk/>; Christensen and Christensen 2007) and ENSEMBLES (<http://ensembles-eu.metoffice.com/>; van der Linden and Mitchell 2009) produced coordinated multi-model ensembles, improving rapidly the knowledge about uncertainties in regional climate projections in the last decade. Furthermore, these simulations provided the basis for most investigations of regional climate change impacts over Europe in recent years.

Due to limited computational capacities, only a limited number of RCM simulations can be realised and it is a question of the experimental design which uncertainty components are primarily tackled within the ensemble. Therefore, missing realisations within RCM ensembles are a common problem and even simple ensemble estimates such as mean and variability are potentially biased due to unequal sampling of the uncertainty components. In order to avoid such biases, Déqué et al. (2007) introduced an iterative data reconstruction method which assumes additivity between uncertainty components in order to estimate the missing climate change signals (CCSs). This reconstruction method was further applied in several studies in order to obtain a balanced design for the analysis of variance components (Déqué et al. 2007; Heinrich and Gobiet 2011b; Prein et al. 2011; Déqué et al. 2012). However, as the method relies on an implicit formulation of the uncertainty components, it cannot be used to extend the ensemble to experiments outside of the original experimental design (e.g., for GCMs that have not been used as driver for any RCM in the ensemble). For such an extension, scaling techniques are widely applied (Mitchell et al. 1999; Mitchell 2003; Rummukainen et al. 2003; Harris et al. 2006; Hingray et al. 2007; Ruosteenoja et al. 2007). They have been originally used to derive regional climate projections of time horizons or emission scenarios which have not been GCM simulated, by scaling the global mean temperature change of simple energy balance models with the geographical pattern of the GCM simulations (pattern-scaling). In our case, we aim at predicting the RCM response from the driving GCM response, by applying a scaling relationship according to the regional CCSs. In this respect, it has already been shown that scaling has skill at the GCM grid point scale (Kendon et al. 2010).

The aim of our study is to assess and compensate for the potential GCM sampling bias in expected regional mean climate change and the associated uncertainty of the ENSEMBLES RCM projections by data reconstruction and combination with the much larger GCM ensemble of the

third phase of the Coupled Model Intercomparison Project (CMIP3; Meehl et al. 2007). For this purpose, we first assess the skill of different statistical additive and scaling reconstruction methods in reproducing ensemble mean and standard deviation. We then apply the most appropriate reconstruction method to the sparsely filled ENSEMBLES simulation matrix and further extend the matrix to all available GCM simulations of the CMIP3 ensemble forced by the A1B emission scenario (Nakicenovic et al. 2000). As final result, we present an update of expected regional climate change for air temperature and precipitation amount in eight European subregions until the mid-twenty-first century and reassess its uncertainty under the light of the extended ensemble.

The paper is structured as follows. Section 2 introduces the data and study regions. In Sect. 3 we describe the reconstruction methods. In Sect. 4 we explain the setup of the cross-validation and the statistical analysis. In Sect. 5 we present the cross-validation results. Section 6 provides a revision of expected regional climate change and its uncertainty over Europe, followed by Sect. 7 which sums up the key findings of this study.

## 2 Climate model data and study region

We use the RCM data from the ENSEMBLES project which produced a set of 21 high resolution RCM simulations with a horizontal grid spacing of about 25 km (Table 1). The ensemble consists of 8 GCMs and 16 RCMs, but due to limited computational resources, only a small fraction (16.4 %) of the possible GCM–RCM combinations could be realised. Sub sampling mainly addressed uncertainty in boundary conditions (choice of the driving GCM) and RCM model formulation (Christensen et al. 2010). Since the choice of the GHG emission scenario is less important until the mid-twenty-first century (Hawkins and Sutton 2009, 2011; Prein et al. 2011), only the A1B emission scenario was used to force the climate simulations. As suggested by Christensen et al. (2010), we consider the three sensitivity experiments of HadCM3–HadRM3 as different model combinations, as their climate response is highly variable (Collins et al. 2006).

Many of the RCM simulations (10 out of 21) were driven by only two GCMs, namely ECHAM5 and HadCM3Q0. All GCMs, except CGCM3.1 and IPSL-CM4, drive at least two RCMs. Three RCMs were forced by multiple GCMs, namely DMI-HIRHAM, METNO-HIRHAM, and SMHIRCA. The driving GCM data of ENSEMBLES was either obtained by the database of ENSEMBLES (<http://ensemblesrt3.dmi.dk/>) or CMIP3 (<https://esg.llnl.gov:8443/>). As we aim at extending the uncertainty analysis to

**Table 1** The ENSEMBLES simulation matrix of the 25 km runs until 2050

		ENSEMBLES GCM Uncertainty										CMIP3 GCM and Internal Uncertainty																		
		app.																												
		cal. / app.																												
		ARPEGE	BCCR-BCM2.0	CGCM3.1 (T47-T4)	ECHAM5 (r3)	HadCM3Q0	HadCM3Q3	HadCM3Q16	IPSL-CM4	CCSM3*	CGCM3.1 (T47)*	CGCM3.1 (T63)	CNRM-CM3	CSIRO-Mk3.0	CSIRO-Mk3.5	ECHAM5*	ECHO-G*	FGOALS-g1.0*	GFDL-CM2.0	GFDL-CM2.1	GISS-AOM*	GISS-EH*	GISS-ER*	INGV-SXG	INM-CM3.0	MIROC3.2 (hi)	MIROC3.2 (med)*	MRI-CGCM2.3.2*	UKMO-HadCM3	UKMO-HadGEM1
ENSEMBLES RCM Uncertainty	C4I-RCA								X																					
	CNRM-RM5.1	X																												
	DMI-HIRHAM	X	X																											
	ETHZ-CLM				X																									
	GKSS-CLM																													
	ICTP-RegCM					X																								
	KNMI-RACMO					X																								
	METNO-HIRHAM		X				X																							
	METOHC-HadRM3Q0						X																							
	METOHC-HadRM3Q3							X																						
	METOHC-HadRM3Q16																													
	MPI-REMO					X																								
	OURANOS-CRCM				X																									
	SMHI-RCA		X		X		X																							
	UCLM-PROMES						X																							
	VMGO-RRCM						X																							

The orange coloured cells marked with X's indicate the available simulations and empty cells represent the missing GCM–RCM combinations. The models spanning the RCM and GCM uncertainty of ENSEMBLES are highlighted in blue and green, respectively. Additional uncertainty due to the CMIP3 GCMs is displayed in red and GCMs which are driven by different perturbed initial conditions are marked with an asterisk. The GCMs and RCMs of ENSEMBLES are used for calibrating the statistical reconstruction methods which are then applied to the GCMs of both ENSEMBLES and CMIP3 in order to reconstruct the associated missing RCM responses

unknown GCMs, we also take into account all additional available GCM simulations of CMIP3 forced by the A1B emission scenario. Altogether, we have 53 GCM simulations for air temperature and 50 for precipitation amount from a set of 27 GCMs (precipitation from 3 out of 5 simulations of the GISS-ER model are missing in the CMIP3 database). From this ensemble, 10 GCMs were started with different perturbed initial conditions and, therefore, cover uncertainty due to natural variability (see Table 1).

In order to be comparable to previous studies conducted within PRUDENCE and ENSEMBLES, we focus on the land grid points of eight European subregions according to Christensen and Christensen (2007): Iberian Peninsula (IP), Mediterranean (MD), France (FR), Middle Europe (ME), Alps (AL), Eastern Europe (EA), British Isles (BI), and Scandinavia (SC). In addition, we also provide information concerning entire Europe (EU). The selected subregions cover the diversity of main climate characteristics in Europe, ranging from arid climate conditions during summer in the southern European regions IP and MD to humid maritime climate characteristics in BI and SC (Heinrich and Gobiet 2011a). The focus is on seasonal mean CCSs, calculated as the difference between the two periods of 2021–2050 for the future period and 1961–1990 for the baseline period (except for the GKSS-CLM model which starts in 1963). The precipitation CCSs are calculated relatively with respect to the baseline period.

### 3 Data reconstruction methods

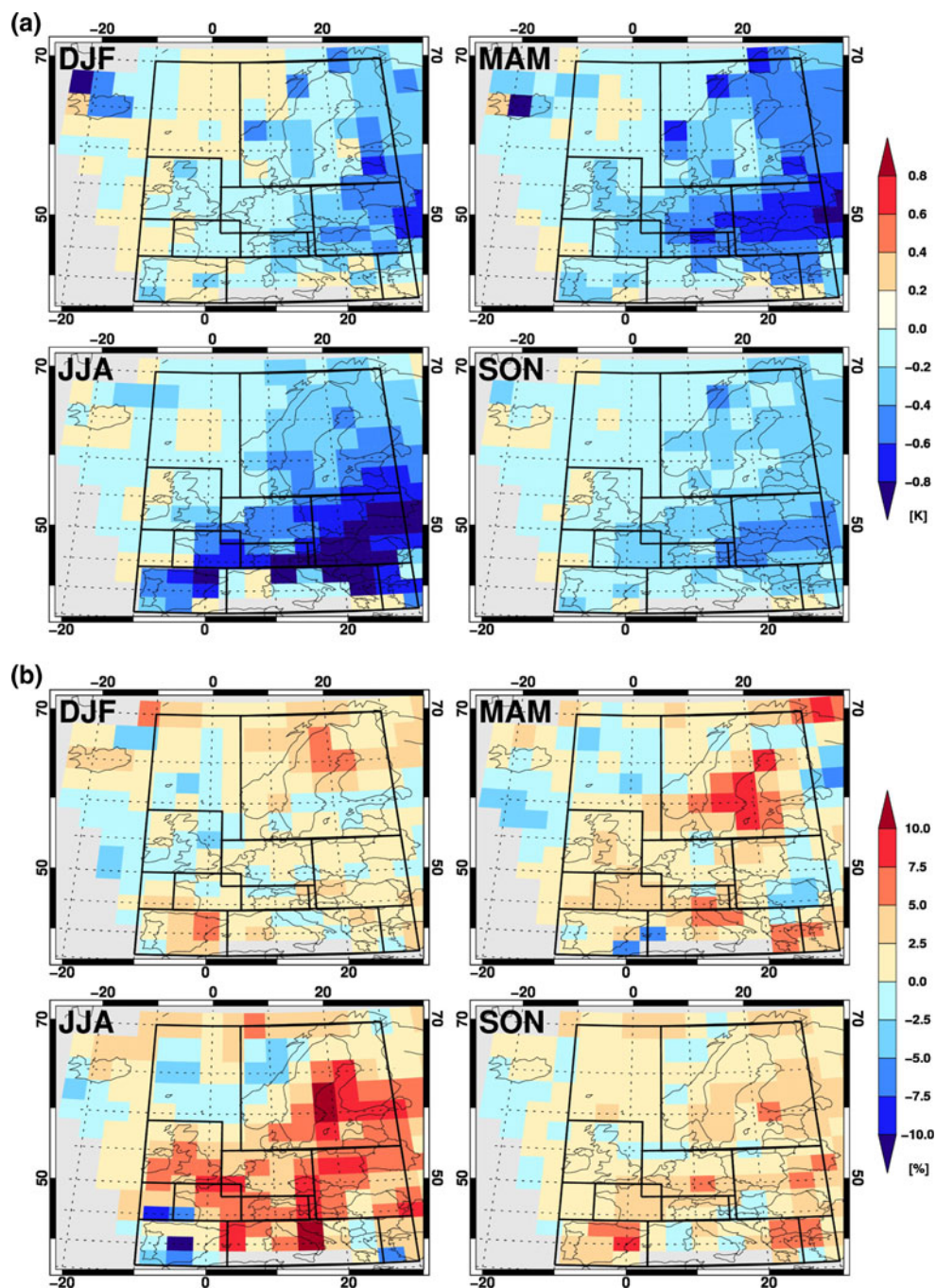
As we aim at deriving RCM responses based on their associated driving GCMs, we first highlight the effect of RCMs on the CCSs of their driving GCMs. Figure 1 shows the mean difference in the seasonal mean CCSs between RCMs and associated driving GCMs of the 21 available ENSEMBLES projections for air temperature and precipitation amount, respectively. The climate model data have been resampled to a common grid of 3.75° longitude and 2.5° latitude which is most common to the ENSEMBLES GCMs. A more detailed description of the resampling technique can be found in Suklitsch et al. (2008).

Concerning air temperature, the RCMs feature smaller CCSs than their driving GCMs for most of Europe with highly pronounced differences in spring and summer, particularly in Eastern and Southern Europe. On subregional scale, EA and MD show the largest difference with −0.8 K in JJA (see Table S1 in the online supplementary material). For precipitation, the RCMs tend to feature larger CCSs (moister conditions) for large parts of Europe, particularly in JJA with the most pronounced differences of +6.5 % in FR, +5.4 % in EA, and +5.1 % in ME.

#### 3.1 Additive method

First, we consider the additive reconstruction method introduced by Déqué et al. 2007 (D07). The reconstruction

**Fig. 1** The mean difference in the seasonal CCSs between RCMs and associated driving GCMs of ENSEMBLES for air temperature (a) and precipitation amount (b). The CCSs are calculated between the two periods of 2021–2050 and 1961–1990. In each panel, *top-left* is winter (DJF), *top-right* is spring (MAM), *bottom-left* is summer (JJA), and *bottom-right* is autumn (SON)



method is embedded in the framework of an analysis of variance (ANOVA), neglecting the highest interaction term in order to reconstruct the actual missing value. The reconstruction algorithm writes in case of the ENSEMBLES simulation matrix as follows:

$$\begin{aligned} \Delta X_{ij} &= \Delta X_{i\bullet} + \Delta X_{\bullet j} - \Delta X_{\bullet\bullet} \\ &= \Delta X_{\bullet\bullet} + (\Delta X_{i\bullet} - \Delta X_{\bullet\bullet}) + (\Delta X_{\bullet j} - \Delta X_{\bullet\bullet}), \end{aligned} \quad (1)$$

where  $\Delta X$  denotes the CCS of an RCM for a specific subregion,  $i$  is the index of the RCM ( $i = 1, \dots, 17$ ), and

$j$  the index of the driving GCM ( $j = 1, \dots, 8$ ). The dot operator denotes averaging across the corresponding indices. The reconstruction algorithm can be understood intuitively: consider RCM1 driven by a set of GCMs and RCM2 driven by the same GCMs except one. This missing value is then reconstructed by adding the mean difference between RCM2 and RCM1 to RCM1. Since the reconstruction of the missing values depends on the grand mean of the entire simulation matrix, 30 iterations are performed (Déqué et al. 2007).

### 3.2 Scaling methods

The main disadvantage of the D07 method is that it cannot be used for an extension to unknown GCMs, as there is no explicit formulation between the CCSs of RCMs and their associated driving GCMs. For such an extension, scaling techniques are widely applied and in this study we assess the capability of different scaling methods in reconstructing the RCM projections.

Scaling can be generally formulated as:

$$\Delta X_{ij} = \Delta Y_j \cdot \tilde{k} + \tilde{d} + \varepsilon_{ij}, \quad (2)$$

where  $\tilde{k}$  and  $\tilde{d}$  are the linear regression coefficients of a least-square-fit to the data. In order to account for the downscaling effect which is specific for a particular RCM, we introduce the additive term  $\varepsilon_{ij}$ . In this study, three different formulations of the RCM specific downscaling effect  $\varepsilon_{ij}$  are investigated: (1) there exists no RCM specific downscaling effect with  $\varepsilon_{ij} = 0$  (SCA1), (2) the RCM specific downscaling effect is the mean RCM specific residual of the least-square-fit (SCA2), (3) the RCM specific downscaling effect is randomly sampled from a normal distribution with zero mean and standard deviation estimated from the regression residuals of all RCMs (SCA3).

In literature, the regression line is typically forced to cross the abscissa at zero (intercept  $\tilde{d} = 0$  in Eq. 2) which is argued to be reasonable if global mean temperature change is used as predictor for the regional climate response (e.g., Hingray et al. 2007; Ruosteenoja et al. 2007). However, the relevance of this assumption has not been assessed so far for a scaling relationship at the GCM scale and, therefore, we also consider a scaling relationship without intercept (SCA0). SCA0 can be regarded as degenerated linear regression with zero intercept, while scaling (Eq. 2) is a full linear regression with varying intercept and slope. We also note that RCM specific random sampling is not feasible as the majority of RCMs is only driven by a single GCM (see Table 1). In addition, the insufficient sample size at the RCM specific level does not allow for applying a hierarchical linear model (e.g., Gelman and Hill 2009) in order to estimate RCM specific regression lines.

Furthermore, we note that the random sampling approach SCA3 is embedded in the framework of multiple imputation (MI; Rubin 1987; Little and Rubin 2002), which aims at generating a set of multiple plausibly reconstructed data sets. Common statistical analysis is then applied to each individual data set, generating a set of parameter estimates of interest. The final parameter estimate is then achieved by simply taking the average of the individual realizations (Rubin 1987). As

we are dealing with a large fraction of missing values, we generate a set of 1,000 reconstructed data sets throughout the study in order to ensure convergence of random sampling.

## 4 Design of the cross-validation and statistical analysis

### 4.1 Cross-validation

In order to compare the skill of the different reconstruction methods, we follow a cross-validation (CV) strategy. CV is often used to estimate the predictive skill of statistical models in application and has the further advantage that it can be applied to small sample sizes as it does not rely on asymptotic theory. In this study, we apply a leave-one-out cross-validation (LOOCV) method by fitting the statistical model to a training data set which consists of leaving out a data point from the original sample. This data point is then predicted by the statistical model and the procedure is repeated for each single data point in the sample, generating a completely reconstructed data set. As the focus of our study is on the estimation of expected seasonal mean climate change and its uncertainty, the LOOCV is used to assess the predictive skill of the reconstruction methods in estimating ensemble mean and standard deviation by comparing their match between original and reconstructed sample. Although an unbiased estimation of the ensemble mean is an important prerequisite for the further analysis of this study, the mean errors might compensate among the predictions and, therefore, the root mean square deviations (RMSD) between original and reconstructed CCSs are also assessed.

Cross-validation (CV) of the methods which are based on RCM specific information (D07 and SCA2) requires RCMs which are driven by more than one GCM. Unfortunately, only a subset of eight simulations of three RCMs were forced by multiple GCMs in ENSEMBLES (see Table 1) and the corresponding LOOCV results have to be regarded as a rather rough estimate of the skill due to the small sample size. However, the LOOCV based on this subset is the only way to compare the skill of all implemented reconstruction methods. All other reconstruction methods allow for a LOOCV based on the full sample of 21 simulations.

Concerning the LOOCV, we follow two strategies. First, we compute a LOOCV which is based on leaving out the CCS of a single RCM (LOOCV-RCM). As most of the GCMs are used as driving data for at least two RCMs, the entire GCM information is available for predicting the RCM change in most cases. Therefore, this strategy can be regarded as measuring the performance of deriving the missing CCSs of the ENSEMBLES simulation matrix (see

Table 1). However, it might be argued that the influence of a single data point decreases as the sample size increases. In order to take this into account and as we also aim at extending the ENSEMBLES simulation matrix to unknown GCMs, we additionally apply a more stringent cross-validation which is achieved by leaving out the CCSs of a driving GCM and all associated RCMs as second LOOCV strategy (LOOCV-GCM). Compared to the LOOCV-RCM, the sample size of the training data sets for each of the LOOCV-GCM predictions is generally smaller. Furthermore, we note that the LOOCV-GCM is asymmetric in the sense that the number of neglected RCM simulations is GCM-dependent (up to almost 25 % of the data is neglected in case of ECHAM5 and HadCM3Q0) which consequently increases the independence among the training samples since the entire information of the driving GCM is removed for each of the RCM predictions.

## 4.2 Statistical analysis

### 4.2.1 Statistical significance

After reconstructing the missing values of the simulation matrix, we assess the statistical significance of the differences in expected seasonal mean climate change and the associated uncertainty between original and reconstructed ensemble. Statistical tests for differences in the mean generally require equal variances (Behrens–Fisher problem). However, this cannot be assumed here, as the reconstruction of the missing CCSs potentially changes the spread of the ensemble. Therefore, we apply a *t* test for unequal variances (Welch-test) under the null hypothesis that the ensemble mean of original and reconstructed ensemble are equal. The application of the *t* test requires independently and normally distributed samples. As we are dealing with rather small sample sizes, it is difficult to assess the normality of the data. However, the assumption of normality in our study is supported by the central limit theorem since multiple averaged quantities are applied. The statistical significance of the differences in the ensemble spread is assessed by applying the robust Fligner–Killeen test (Conover et al. 1981) under the null hypothesis that the ensemble variances of original and reconstructed ensemble are equal.

In order to eliminate possible dependencies among the RCM simulations due to their driving GCMs, we average across the RCMs according to their driving GCMs. For the same reason, we average across the different GCM runs of the CMIP3 database after reconstructing the RCM changes. Therefore, the sample sizes of original and reconstructed simulation matrices reduce to the number of GCMs indicated in Table 1 (8 for the ENSEMBLES simulation matrix and 27 for the extended matrix).

Significance levels lower than 3, 3 to 5, 5 to 10 %, and greater than 10 % are termed as strongly significant, significant, weakly significant, and insignificant, respectively.

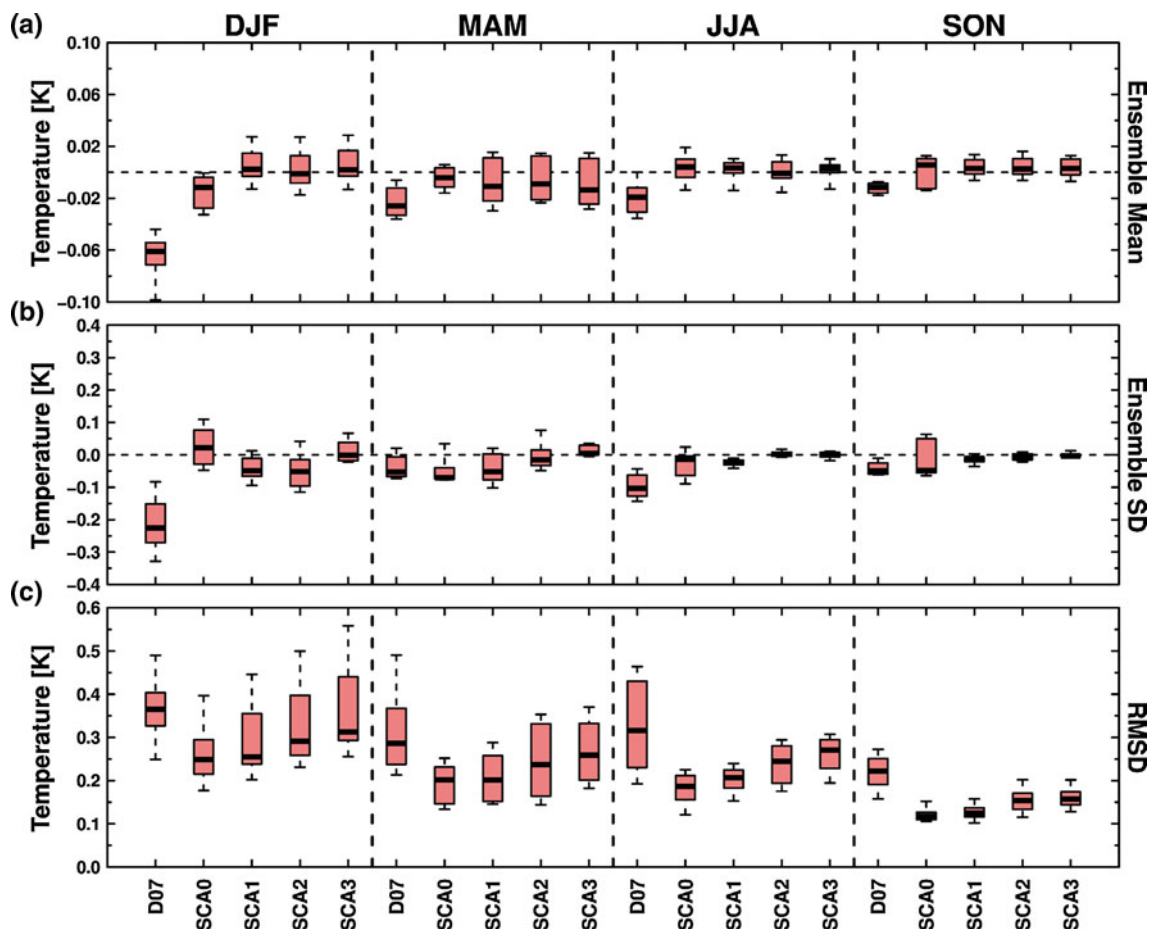
### 4.2.2 Quantification of uncertainty

The uncertainties of the projected changes are quantified by two measures. Firstly, we calculate the ensemble standard deviation. As its calculation is based on rather small sample sizes in this study, we apply a minor bias correction as proposed in Knutti et al. (2010). Secondly, we calculate the percentage of models which coincide in the sign of change as a non-parametric uncertainty measure. Applying the confidence terminology defined by the intergovernmental panel on climate change (Solomon et al. 2007; Mastrandrea et al. 2010), very high confidence, high confidence, and medium confidence is obtained if at least 90 %, at least 80 %, and at least 50 % agree in the sign of the multi-model mean change, respectively.

## 5 Results of the cross-validation

Figures 2 and 3 depict the results of the LOOCV for air temperature and precipitation amount, respectively. The spread of the box-whisker plots displays variability among subregions. For comparative purposes, we only show the results of LOOCV-RCM based on the subset of RCM simulations driven by multiple GCMs which includes all reconstruction methods. All additional LOOCV results based on LOOCV-GCM and the full sample are provided in the online supplementary material, Fig. S1 to Fig. S6.

For air temperature, the largest differences between reconstructed and original ensemble mean are obtained for D07 concerning all seasons. However, the bias is rather small, showing mostly an underestimation of the ensemble mean. The largest bias of D07 is obtained in DJF with  $-0.06$  K in the median and ranging from  $-0.13$  K to  $-0.04$  K. SCA0 shares this negative bias in DJF, but with a lower magnitude of  $-0.01$  K in the median. For all other methods, biases in both directions are obtained with magnitudes less than 0.02 K in the median. The differences between SCA1 and SCA3 in the ensemble mean can be related to finite sampling of the residuals and converge to zero as the number of random draws increases. For the ensemble standard deviation, D07 generally shows an underestimation which is largest in DJF with  $-0.23$  K in the median. Although the scaling methods generally perform better in reconstructing the ensemble standard deviation, underestimated variability is obtained as expected for the scaling relationships without random sampling of the residuals. Only the randomized scaling approach SCA3 is able to compensate for the underestimated standard



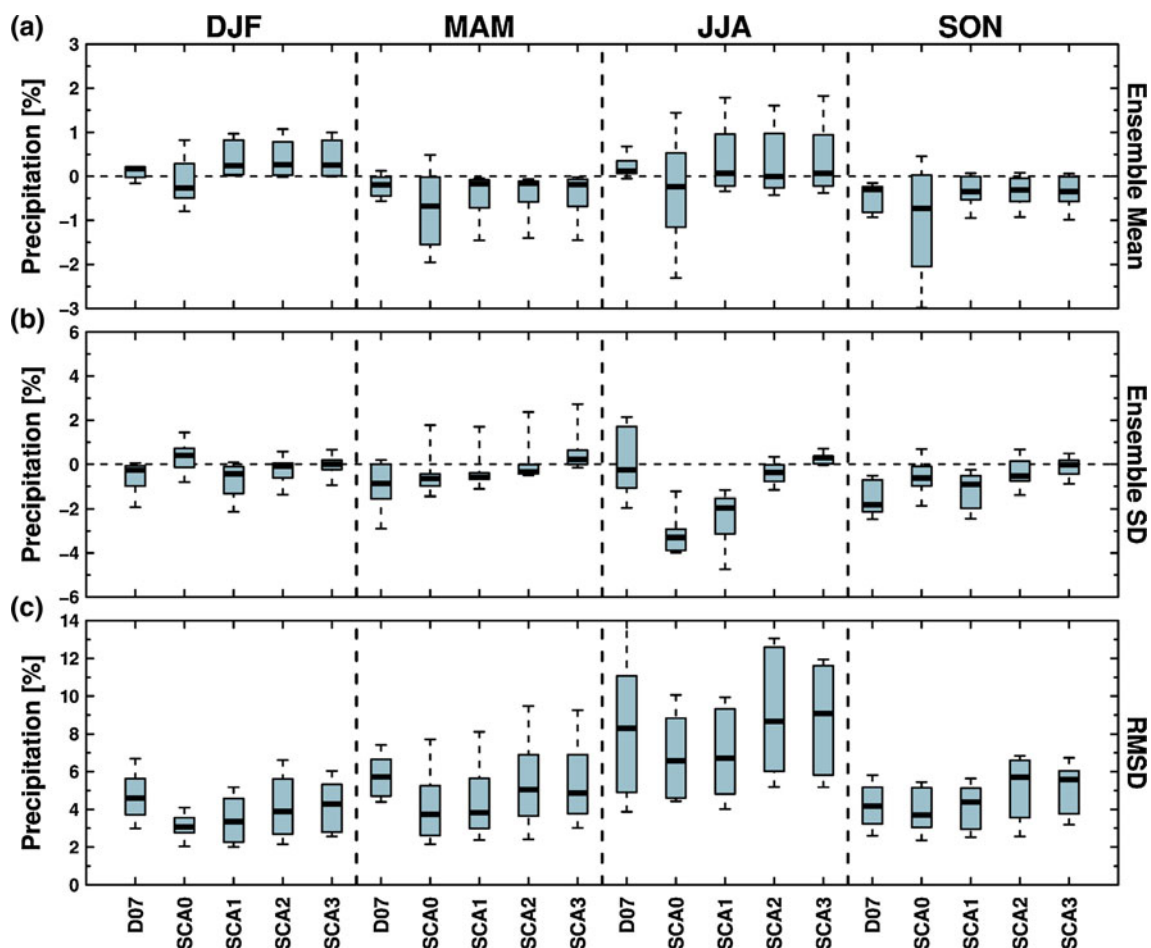
**Fig. 2** Results of the LOOCV-RCM for air temperature based on the subset of RCMs which are driven by multiple GCMs. Displayed are the differences between reconstructed and original ensemble mean (a) and standard deviation (b), and the associated root mean square

deviations (c). The spread of the *box-whisker plots* shows variability among subregions and displayed are the 10th, 25th, 50th, 75th, and 90th percentile

deviation as additional variability is introduced due to random sampling of the residuals. The RMSD is generally lower for the scaling methods. The difference in the RMSD between SCA1 and SCA3 can be explained by the fact that additional variability is introduced due to random sampling of the residuals. More specifically, the difference in the squared RMSD between SCA1 and SCA3 can be calculated analytically and is given by the mean variance of the residuals among the LOOCV samples. All reconstruction methods show a seasonal dependency of the RMSD with the best and worst skill in DJF and SON, respectively. The differences between the various scaling methods are generally small (especially for SCA0 and SCA1).

For precipitation amount, D07 tends to overestimate (underestimate) the ensemble mean in DJF and JJA (MAM and SON). The biases are rather small with a peak magnitude in SON with  $-0.3\%$  in the median. For the scaling methods, the ensemble mean is underestimated in MAM and SON. SCA0 generally shows large biases in combination with rather large ranges among the subregions. The

largest bias of SCA0 is found in SON with a median difference of  $-0.7\%$  and ranging from  $-3.0\%$  to  $+1.0\%$ . For the ensemble standard deviation, D07 again tends to underestimate the ensemble spread with the largest difference in the median of  $-1.8\%$  in SON. The scaling methods partly show a pronounced underestimation of variability. The largest difference in the median is obtained for SCA0 and SCA1 in JJA with a magnitude of  $-3.3\%$  and  $-2.0\%$ , respectively. As already shown for air temperature, SCA3 again shows the overall best performance in reproducing the ensemble spread as random sampling of the residuals increases variability of the scaling relationship and consequently compensates for the underestimated ensemble standard deviation. For the RMSD, the scaling reconstruction methods consistently reveal the best skill in DJF while D07 shows the lowest RMSD in SON. The differences in the RMSD between SCA0 and SCA1 are again small. The worst skill for all reconstruction methods is obtained in JJA, peaking up to  $8.3\%$  for D07. Here, the rather large difference between SCA1 and SCA3 can be



**Fig. 3** Same as Fig. 2 but for precipitation amount

related to large additional variability introduced by random sampling of the residuals and reasonably indicates large variability of the summer precipitation CCSs among the different RCM projections (e.g., Déqué et al. 2012).

The results of the second cross-validation strategy LOOCV-GCM, which is based on leaving out the CCSs of a driving GCM and all associated RCMs, underpin the results reported above. Extending the LOOCV to the entire sample, rather large negative precipitation biases are obtained for SCA0 in all seasons. Concerning the ensemble spread, SCA0 and SCA1 mostly underestimate variability and the superiority of SCA3 in reconstructing the ensemble spread is further confirmed (see online supplementary material, Fig. S1 to Fig. S6).

Based on the results of the LOOCV, we choose SCA3 as preferred reconstruction method, especially due to its overall best performance in reconstructing ensemble mean and variability. Table 2 summarises the results of the two LOOCV strategies for SCA3 based on the entire sample as further applied in the study. As expected, the results of the more stringent LOOCV-GCM strategy indicate a worse performance than the LOOCV-RCM strategy which can be

related to the smaller size of the training samples and the asymmetric character of LOOCV-GCM. However, both CV strategies reveal differences in the ensemble mean and spread mostly clearly less than 0.1 K and 1.0 % for air temperature and precipitation amount, respectively. As these values provide a measure of the predictive skill of our statistical reconstruction method, we expect rather small errors introduced by the statistical model in prediction mode. As already discussed above, the RMSD values are inflated by additional variability due to random sampling of the residuals. The coefficient of determination ( $R^2$ ) is generally larger for air temperature than for precipitation amount for MD in JJA up to 0.9 for air temperature in several subregions and seasons. However, small  $R^2$  values are generally not related to large biases and/or differences in the ensemble spread, reflecting a potential instability of the  $R^2$  due to the rather small sample size. Figure 4 displays the linear fits between the CCSs of RCMs and associated driving GCMs as further applied in the study. As it can be seen, intercept and slope are both varying across subregions and seasons, further underpinning that the



formulation of the zero intercept model (SCA0) may not be appropriate. This is generally in agreement with the results of the LOOCV. For example, air temperature reveals a distinct non-zero intercept in MAM for all subregions which consequently results in poor performance of SCA0 concerning both LOOCV strategies (see online supplementary material, Fig. S3 and Fig. S5).

## 6 Revision of expected regional climate change and its uncertainty over Europe

Table 3a and b show the differences in ensemble mean and standard deviation between original and filled ENSEMBLES simulation matrix for air temperature and precipitation amount, respectively. For air temperature, the differences in the ensemble mean are mostly negative in DJF, JJA, and SON, indicating slightly larger CCSs of the original ensemble, while in MAM larger mean values of the filled ensemble are obtained for all subregions. However, the differences are rather small and insignificant, with maximum differences in the order of  $\pm 0.2$  K. For the ensemble standard deviation, mostly lower values for the filled ensemble are obtained. Exceptions are EA in JJA and BI in DJF and the differences are again small and insignificant, peaking up to  $-0.5$  K for SC in MAM. For precipitation amount, the differences in the ensemble mean vary across seasons and subregions. As for air temperature, the differences are again small and insignificant, with a maximum difference of  $+1.1$  % for AL in DJF. Concerning the ensemble spread, mostly lower standard deviations for the filled ensemble are obtained. Exceptions are SC in DJF as well as FR and ME in MAM. The differences are again rather small and insignificant, peaking up to  $-4.4$  % for FR in JJA.

Table 3c and d show the differences in ensemble mean and standard deviation between original and extended ensemble concerning all available GCM simulations of CMIP3 forced by the A1B emission scenario. For air temperature, the ensemble mean of the extended ensemble is mostly reduced, showing the largest differences in EA in JJA with  $-0.4$  K. However, no statistical significance for the changes is obtained. Concerning the ensemble spread, the extended ensemble generally shows a reduction with the most pronounced changes in EA and SC in MAM with  $-0.5$  K. Again, the changes are mostly insignificant. For precipitation amount, the sign of the difference in the ensemble mean is varying among subregions and seasons (only in SON mostly lower CCSs of the extended ensemble are obtained). The largest difference is  $-4.2$  % in IP in DJF and the changes are again mostly insignificant. Concerning the ensemble spread, mostly lower ensemble standard deviations for the extended ensemble are

obtained. The most pronounced change is obtained in AL in JJA with  $-3.3$  % and the changes are again mostly insignificant.

We note that the main characteristics of the geographical pattern of the CCSs are not deteriorated through statistical reconstruction (also see online supplementary material, Fig. S7 and Fig. S8). These patterns include large spatial differences of winter air temperature change, which are explained by moderate warming of the ocean influencing the maritime climate of western Europe in combination with altered snow-albedo feedback mechanisms in northern and eastern Europe (Rowell 2005), high summer air temperatures in the south, which are related to an earlier and more rapid reduction of soil moisture in spring (e.g., Wetherald and Manabe 1995; Gregory et al. 1997), and a dipolar pattern of the precipitation change (decrease in the south and increase in the north), which can be related to a seasonal dependent northward shift of the mid-latitude storm track which is identified as the European Climate change Oscillation (ECO) (Giorgi and Coppola 2007). Although the focus of this study is on near-term projections until the mid-twentyfirst century, we stress that the projected changes are expected to further increase after the 2050s (e.g., Heinrich et al. 2013). In this respect, the projected precipitation changes are expected to further intensify until the end of the twentyfirst century with the most pronounced intensifications for the hot-spot regions in southern and northern Europe. Furthermore, the confidence of the projected precipitation changes generally increases until the end of the twentyfirst century. However, low confidence remains especially along the transition zone from drier conditions in southern Europe to wetter conditions in northern Europe.

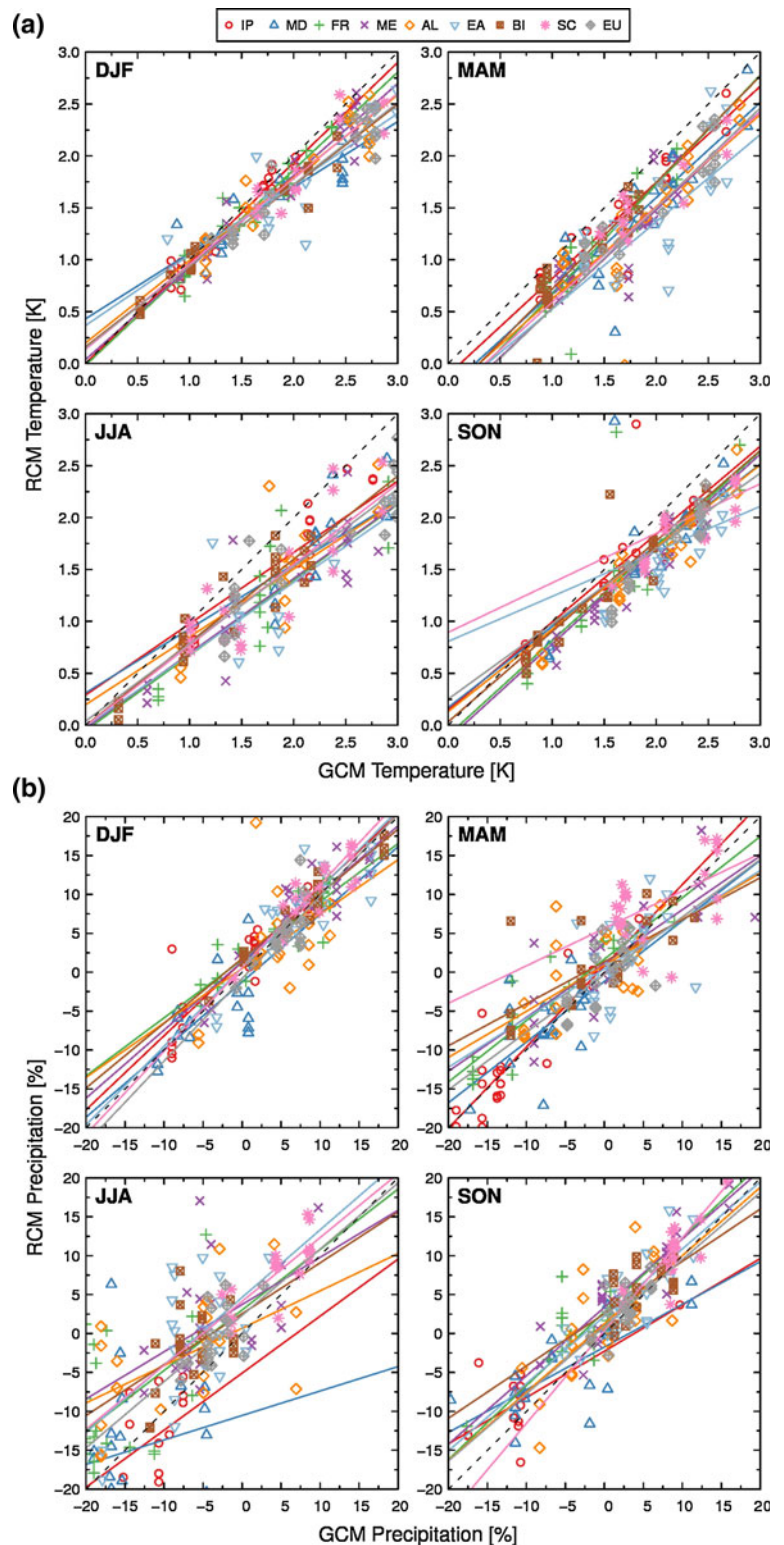
Figure 5 depicts maps for the confidence levels of the precipitation changes for the original, filled, and extended ENSEMBLES simulation matrix. Maps for air temperature are not shown, since very high confidence of warming is achieved in all subregions with and without reconstruction (see online supplementary material, Fig. S9). The confidence of the projected precipitation changes are partly reduced for the filled ENSEMBLES simulation matrix in DJF, MAM, and SON, revealing overconfident projections of the original ensemble. Concerning the extended ENSEMBLES simulation matrix, both lower and larger confidence levels are obtained, revealing under- as well as overconfident projections of the original ensemble. For example, larger confidence levels are obtained for the changes in MD in SON, while in BI, EA, IP, and ME the confidence is reduced. However, the differences between original, filled, and extended ENSEMBLES ensemble are generally small and the overall picture of high confidence in the northern- and southernmost European regions with low confidence in-between is further underpinned.

**Table 2** Results of the cross-validation for SCA3

	Air temperature				Precipitation amount			
	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
<b>IP</b>								
Mean	-0.1/0.2	-0.3/-3.8	0.5/3.4	0.2/2.9	0.1/2.8	-1.1/-4.3	-1.8/28.3	2.5/-15.4
SD	0.1/-0.1	-0.4/1.3	-0.5/-3.3	0.1/-0.6	-0.1/-18.4	-3.0/11.1	-0.1/-20.7	-7.2/16.8
RMSD	0.2/0.2	0.3/0.3	0.3/0.4	0.4/0.5	4.5/4.2	5.1/5.0	10.5/10.4	5.7/5.8
R <sup>2</sup>	0.9/0.9	0.9/0.9	0.9/0.8	0.8/0.8	0.8/0.8	0.8/0.8	0.3/0.3	0.5/0.6
<b>MD</b>								
Mean	0.0/3.2	-0.7/-9.2	1.1/5.0	0.9/9.0	-8.8/-1.1	8.8/-9.3	-3.5/-0.5	-2.0/1.2
SD	0.1/0.4	-0.3/8.3	-0.2/1.0	-1.2/-7.2	-6.9/-0.4	-14.6/9.2	2.2/-9.6	3.9/-20.2
RMSD	0.3/0.3	0.5/0.5	0.5/0.5	0.6/0.6	5.4/5.5	6.5/7.3	11.8/11.8	5.4/5.5
R <sup>2</sup>	0.8/0.8	0.7/0.7	0.8/0.7	0.5/0.5	0.6/0.6	0.4/0.4	0.0/0.1	0.6/0.6
<b>FR</b>								
Mean	0.0/1.5	-0.1/-3.5	0.4/2.9	0.2/2.7	-0.5/0.4	6.7/45.2	-13.3/-7.3	24.2/84.4
SD	-0.3/-1.3	0.1/2.2	-0.7/-3.6	-0.7/0.0	-0.9/13.0	4.7/-38.4	-7.2/15.7	-25.5/-123.7
RMSD	0.2/0.2	0.3/0.3	0.5/0.5	0.5/0.5	3.4/3.3	4.3/4.5	11.3/11.3	6.2/7.2
R <sup>2</sup>	0.9/0.9	0.8/0.8	0.8/0.8	0.7/0.7	0.8/0.8	0.8/0.8	0.2/0.2	0.7/0.7
<b>ME</b>								
Mean	0.1/3.7	-0.9/-13.1	0.4/1.2	0.0/3.3	12.0/33.9	7.7/18.4	0.5/22.3	-2.9/-10.9
SD	0.0/0.6	0.2/10.5	-0.6/-2.9	-0.2/0.4	-5.3/1.5	12.0/-10.3	-11.7/11.8	4.4/28.5
RMSD	0.3/0.3	0.6/0.7	0.4/0.4	0.6/0.6	4.1/4.3	6.6/6.4	8.8/8.9	4.4/4.5
R <sup>2</sup>	0.9/0.9	0.5/0.5	0.8/0.7	0.5/0.5	0.8/0.8	0.7/0.6	0.3/0.3	0.8/0.8
<b>AL</b>								
Mean	0.0/5.8	-0.7/-14.4	0.7/3.2	0.2/6.6	19.6/48.3	4.3/68.9	3.9/62.8	5.8/12.7
SD	0.0/3.1	0.3/14.5	-0.1/-4.2	-1.2/-3.0	-18.2/-22.4	-5.4/-47.1	13.0/20.8	5.3/-6.1
RMSD	0.3/0.3	0.5/0.6	0.5/0.5	0.7/0.7	7.6/8.0	6.2/6.7	9.4/9.4	7.0/7.2
R <sup>2</sup>	0.9/0.9	0.6/0.6	0.8/0.7	0.5/0.4	0.3/0.3	0.4/0.4	0.3/0.3	0.5/0.5
<b>EA</b>								
Mean	-0.5/3.7	-0.8/-13.8	-0.4/1.5	0.0/6.1	9.9/78.8	13.1/55.6	23.3/80.1	-2.1/-9.7
SD	0.4/3.3	0.6/12.3	0.5/-1.1	1.1/3.7	-14.8/15.7	31.0/9.4	-19.7/-36.0	2.1/20.8
RMSD	0.4/0.4	0.8/0.9	0.6/0.6	0.8/0.8	6.6/6.6	7.3/7.6	9.4/10.2	4.8/4.8
R <sup>2</sup>	0.8/0.8	0.5/0.5	0.7/0.7	0.1/0.1	0.7/0.7	0.2/0.2	0.3/0.4	0.6/0.6
<b>BI</b>								
Mean	0.1/2.0	0.1/-1.4	0.4/0.9	0.1/1.3	4.2/6.2	-4.4/-54.8	13.5/40.7	1.5/34.5
SD	0.0/-1.0	0.1/0.6	-0.4/-2.0	-0.2/2.2	0.5/26.4	0.1/51.7	-2.0/50.6	-0.9/29.7
RMSD	0.2/0.2	0.3/0.3	0.3/0.3	0.3/0.3	2.0/2.1	5.8/5.8	6.3/7.0	3.7/3.9
R <sup>2</sup>	1.0/1.0	0.8/0.8	0.9/0.9	0.8/0.8	0.9/0.9	0.5/0.5	0.2/0.3	0.4/0.5
<b>SC</b>								
Mean	-0.5/0.2	0.8/1.9	-0.3/-3.8	-0.5/3.6	0.1/-19.8	-8.3/-57.9	0.9/42.1	-10.5/-49.1
SD	-0.2/2.4	0.1/11.8	0.3/1.6	0.2/1.3	-5.1/-37.5	-2.8/-10.9	1.4/-32.5	0.6/63.8
RMSD	0.4/0.4	0.8/0.8	0.4/0.4	0.7/0.7	4.1/4.1	6.2/6.3	3.7/3.8	4.0/4.5
R <sup>2</sup>	0.9/0.9	0.7/0.6	0.8/0.8	0.2/0.2	0.7/0.7	0.3/0.2	0.6/0.6	0.7/0.7
<b>EU</b>								
Mean	0.0/2.6	-0.6/-7.1	-0.1/0.1	0.0/4.1	2.4/15.9	11.2/76.9	1.0/13.2	-2.2/8.1
SD	0.1/0.8	-0.2/6.9	-0.2/-1.1	-0.2/1.6	-3.9/-10.5	9.7/-33.9	-2.8/4.3	-2.2/-8.5
RMSD	0.2/0.2	0.6/0.6	0.4/0.4	0.6/0.7	3.3/3.2	4.4/4.8	4.4/4.4	2.9/3.0
R <sup>2</sup>	0.9/0.9	0.7/0.6	0.8/0.8	0.3/0.3	0.6/0.6	0.4/0.3	0.4/0.4	0.5/0.5

The left and right values indicate the estimates of LOOCV-RCM and LOOCV-GCM, respectively. Differences in the ensemble mean (Mean) and standard deviation (SD) between reconstructed and original ensemble are multiplied by a factor of 100 and the units are [K] and [%] for air temperature and precipitation amount, respectively. The RMSD values are inflated by additional variability due to random sampling of the residuals. R<sup>2</sup> represents the coefficient of determination between the CCSs of RCMs and associated driving GCMs

**Fig. 4** Linear least-square fits between the seasonal CCSs of RCMs and associated driving GCMs for air temperature (a) and precipitation amount (b). In each panel, *top-left* is DJF, *top-right* is MAM, *bottom-left* is JJA, and *bottom-right* is SON



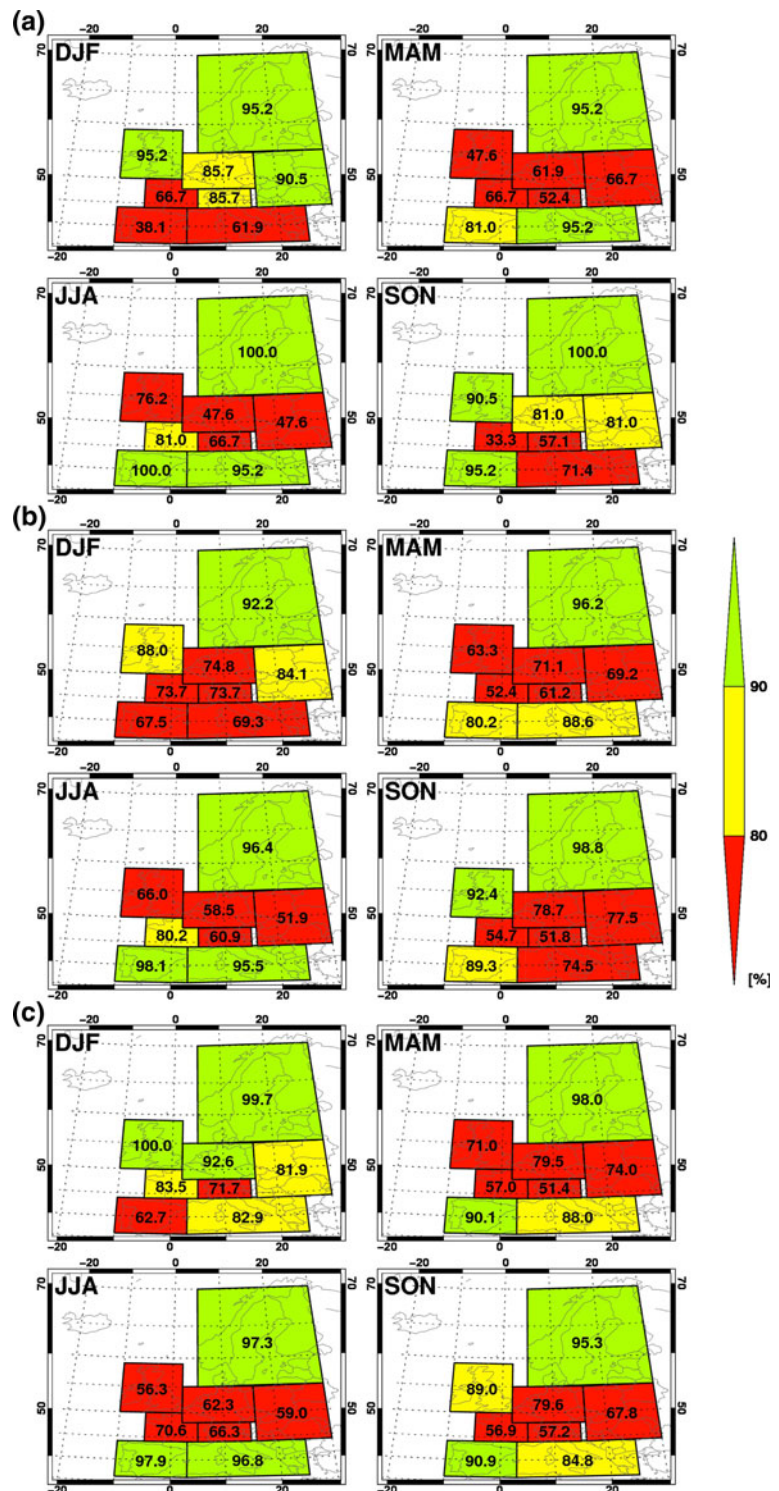
### 7 Summary and conclusions

This study assesses and compensates the GCM sampling bias in expected regional mean climate change and the associated uncertainty of the ENSEMBLES RCM

projections by combining them with the full set of the CMIP3 GCM ensemble. The focus was on eight European subregions and the seasonal mean changes of the two key climate variables air temperature and precipitation amount until the mid-twentyfirst century were assessed.



**Fig. 5** Confidence of the projected precipitation changes between 2021–2050 and 1961–1990 for the original (a), filled (b), and extended (c) ENSEMBLES simulation matrix. Green, yellow, and red colours display very high confidence, high confidence, and medium or no confidence, respectively. The numbers indicate the percentage of models which agree in the sign of the ensemble mean. Percentages lower than 50 % indicate skewed distributions. In each panel, top-left is DJF, top-right is MAM, bottom-left is JJA, and bottom-right is SON



In order to underpin the importance of reconstructing RCM responses, we first highlighted the impact of RCMs on the CCSs of their driving GCMs and showed that down-scaling generally leads to less warming (up to a reduction of 30 % of the GCM CCS) and partly to more precipitation over Europe than projected by GCMs. We introduced and evaluated various statistical data reconstruction methods which

mimic these effects and allow for an extension of the RCM ensemble to additional driving GCMs. The most appropriate reconstruction method, a randomized scaling approach with errors in ensemble mean and standard deviation mostly less than 0.1 K and 1.0 % for air temperature and precipitation amount, respectively, was applied to fill the missing values of the ENSEMBLES simulation matrix and to further extend

the matrix by all available CMIP3 GCM simulations forced by the A1B emission scenario. Differences between original, filled, and extended ENSEMBLES simulation matrix were assessed in order to identify potential ensemble biases and improperly estimated uncertainty ranges due to the GCM sampling strategy of ENSEMBLES. The key findings of the intercomparison can be summarised as follows: (1) The estimated mean climate change for air temperature and precipitation amount over Europe is not significantly altered due to reconstruction in almost all seasons and subregions. (2) The estimated uncertainty is generally not increased by extension of the ENSEMBLES simulation matrix to the entire CMIP3 ensemble. In some seasons and subregions it is even significantly reduced. We note that no weighting regarding model performance was applied in our study. Although our framework would allow for the incorporation of such weights, the benefits of performance weighting in constraining uncertainty of future projections could yet not be sufficiently demonstrated (e.g., Knutti et al. 2010; Déqué and Somot (2010)).

From the results of our analysis we conclude that expected mean climate change and the associated uncertainty of the ENSEMBLES RCM projections are both not underestimated using only few driving GCMs. In contrast, the ensemble spread of the extended ensemble is partly significantly lower than that of the original ensemble. Therefore, this study substantially adds to the reliability of numerous recent climate change impact studies over Europe which use the full range or a carefully selected subset of the ENSEMBLES projections (e.g., Heinrich and Gobiet 2011a; Finger et al. 2012), since it confirms that the ensemble is not significantly biased and the uncertainty is not underestimated due to GCM sub-sampling.

**Acknowledgments** This work was partly funded by the Austrian Climate Research Program (ACRP) projects reclip: century (A760437; [http://reclip.ait.ac.at/reclip\\_century/](http://reclip.ait.ac.at/reclip_century/)) and El.Adapt (B060380), and by the EU-FP7 project ACQWA (contract number 212250; <http://acqwa.ch/>) whose support is gratefully acknowledged. The ENSEMBLES data used in this work was funded by the EU-FP6 integrated project ENSEMBLES (contract number 505539) whose support is gratefully acknowledged. Furthermore we acknowledge the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and the WCRP's Working Group on Coupled Modelling (WGCM) for making available the WCRP CMIP3 multi-model dataset.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

- Buonomo E, Jones R, Huntingford IC, Hannaford J (2007) On the robustness of changes in extreme precipitation over Europe from two high resolution climate change simulations. *Q J R Meteorol Soc* 133:65–81
- Castro CL, Pielke RA Sr, Leoncini G (2005) Dynamical downscaling: assessment of value retained and added using the regional atmospheric modeling system (rams). *J Geophys Res* 110:1–21
- Christensen JH, Christensen OB (2007) A summary of the Prudence model projections of changes in European climate by the end of this century. *Clim Change* 81:7–30. doi:10.1007/s10584-006-9210-7
- Christensen JH, Rummukainen M, Lenderink G (2010) Formulation of very-high-resolution regional climate model ensembles for Europe. In: van der Linden P, Mitchell JFB (eds) ENSEMBLES: climate change and its impacts: summary of research and results from the ENSEMBLES project. Met Office Hadley Centre, Exeter, UK, pp 47–58
- Collins M (2007) Ensembles and probabilities: a new era in the prediction of climate change. *Phil Trans R Soc A* 365:1957–1970. doi:10.1098/rsta.2007.2068
- Collins M, Booth BBB, Harris GR, Murphy JM, Sexton DMH, Webb MJ (2006) Towards quantifying uncertainty in transient climate change. *Clim Dyn* 27:127–147
- Conover WJ, Johnson ME, Johnson MM (1981) A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data. *Technometrics* 23:351–361
- Déqué M, Somot S (2010) Weighted frequency distributions express modelling uncertainties in the ENSEMBLES regional climate experiments. *Clim Res* 44:195–209. doi:10.3354/cr00866
- Déqué M, Rowell DP, Lüthi D, Giorgi F, Christensen JH, Rockel B, Jacob D, Kjellström E, Castro M, Hurk B (2007) An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections. *Clim Change* 81:53–70
- Déqué M, Somot S, Sanchez-Gomez E, Goodess CM, Jacob D, Lenderink G, Christensen OB (2012) The spread amongst ENSEMBLES regional scenarios: regional climate models, driving general circulation models and interannual variability. *Clim Dyn* 38:951–964. doi:10.1007/s00382-011-1053-x
- Feser F, Rockel B, von Storch H, Winterfeldt J, Zahn M (2011) Regional climate models add value to global model data: a review and selected examples. *Bull Am Meteorol Soc* 92:1181–1192. doi:10.1175/2011BAMS3061.1
- Finger D, Heinrich G, Gobiet A, Bauder A (2012) Projections of future water resources and their uncertainty in a glacierized catchment in the Swiss Alps and the subsequent effects on hydropower production during the 21st century. *Water Resour Res* 48. doi:10.1029/2011WR010733
- Gelman A, Hill J (2009) Data Analysis using regression and multilevel/hierarchical models. Cambridge University Press, New York, USA
- 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 regional climate change simulation: a review. *Rev Geophys* 29:191–216. doi:10.1029/90RG02636
- Giorgi F, Mearns LO (1999) Introduction to special section: regional climate modelling revisited. *J Geophys Res* 104:6335–6352. doi:10.1029/98JD02072
- Gregory JM, Mitchell JFB, Brady AJ (1997) Summer drought in northern midlatitudes in a time-dependent CO<sub>2</sub> climate experiment. *J Clim* 10:662–686
- Harris GR, Sexton DMH, Booth BBB, Collins M, Murphy JM, Webb MJ (2006) Frequency distributions of transient regional climate change from perturbed physics ensembles of general circulation model simulations. *Clim Dyn* 27:357–375. doi:10.1007/s00382-006-0142-8

- Hawkins E, Sutton R (2009) The potential to narrow uncertainty in regional climate predictions. *Bull Am Meteorol Soc* 90:p1095. doi:10.1175/2009BAMS2607.1
- Hawkins E, Sutton R (2011) The potential to narrow uncertainty in projections of regional precipitation change. *Clim Dyn* 37:407–418. doi:10.1007/s00382-010-0810-6
- Heinrich G, Gobiet A (2011a) The future of dry and wet spells in Europe: a comprehensive study based on the ENSEMBLES regional climate models. *Int J Climatol*. doi:10.1002/joc.2421
- Heinrich G, Gobiet G (2011b) Expected climate change and its uncertainty in the Alpine region. WEGC Report to ACRP Nr. 02/2011, Wegener Center for Climate and Global Change, University of Graz, Austria
- Heinrich G, Gobiet A, Truhetz H, Mendlik T (2013). Expected climate change and its uncertainty in the Alpine region—extended uncertainty assessment of the reclip: century and ENSEMBLES multi-model dataset. Scientific Report no. 50, Wegener Center for Climate and Global Change, University of Graz, Austria
- Hingray B, Mezghani A, Buishand TA (2007) Development of probability distributions for regional climate change from uncertain global mean warming and an uncertain scaling relationship. *Hydrol Earth Syst Sci* 11:1097–1114. doi:10.5194/hess-11-1097-2007
- Jones RG, Murphy JM, Noguer M (1995) Simulation of climate change over Europe using a nested regional-climate model. part i: assessment of control climate, including sensitivity to location of lateral boundaries. *Q J R Meteorol Soc* 121:1413–1449
- Kendon EJ, Jones RG, Kjellström E, Murphy JM (2010) Using and designing GCM–RCM ensemble regional climate projections. *J Climate* 23:6485–6503. doi:10.1175/2010JCLI3502.1
- Knutti R, Furrer R, Tebaldi C, Cermak J, Meehl GA (2010) Challenges in combining projections from multiple climate models. *J Climate* 23:2739–2758. doi:10.1175/2009JCLI3361.1
- Laprise R (2003) Resolved scales and nonlinear interactions in limited-area models. *J Atmos Sci* 60:768–779
- Little RJA, Rubin DB (2002) *Statistical analysis with missing data*. Wiley, Hoboken, New Jersey, USA
- Mastrandrea MD, Field CB, Stocker TF et al. (2010) Guidance bote for lead authors of the IPCC fifth assessment report on consistent treatment of uncertainties. Intergovernmental Panel on Climate Change (IPCC), CA, USA
- McGregor JL (1997) Regional climate modeling. *Meteorol Atmos Phys* 63:105–117. doi:10.1007/BF01025367
- Meehl GA, Covey C, Delworth T, Latif M, McAvaney B, Mitchell JFB, Stouffer RJ, Taylor KE (2007) The WCRP CMIP3 multimodel dataset: a new era in climate change research. *Bull Am Meteorol Soc* 88:1383–1394
- Mitchell TD (2003) Pattern scaling: an examination of the accuracy of the technique for describing future climates. *Clim Change* 60: 217–242
- Mitchell JFB, Johns TC, Eagles M, Ingram WJ, Davis RA (1999) Towards the construction of climate change scenarios. *Clim Change* 41:547–581
- Nakicenovic N, Alcamo J, Davis G et al (2000) IPCC special report on emissions scenarios. Cambridge University Press, Cambridge, United Kingdom and New York
- Prein A, Gobiet A, Truhetz H (2011) Analysis of uncertainty in large scale climate projections over Europe. *Meteorol Z* 20:383–395. doi:10.1127/0941-2948/2011/0286
- Rowell PR (2005) A scenario of European climate change for the late twenty-first century: seasonal means and interannual variability. *Clim Dyn* 25:837–849. doi:10.1007/s00382-005-0068-6
- Rubin DB (1987) *Multiple imputation for nonresponse in surveys*. Wiley, USA
- Rummukainen M (2010) State-of-the-art with regional climate models. *WIREs Clim Change* 1:82–96. doi:10.1002/wcc.8
- Rummukainen M, Räisänen J, Bjørge D, Christensen JH, Christensen OB, Iversen T, Jylhä K, Ólafsson H, Tuomenvirta H (2003) Regional climate scenarios for use in Nordic water resources studies. *Nordic Hydrol* 34:399–412
- Ruosteenoja K, Tuomenvirta H, Jylhä K (2007) GCM-based regional temperature and precipitation change estimates for Europe under four SRES scenarios applying a super-ensemble pattern-scaling method. *Clim Change* 81:193–208
- Solomon S, Qin D, Maning M et al. (2007) In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds.) *Climate change 2007: The physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change, technical summary*, Cambridge University Press, Cambridge, United Kingdom and New York, USA
- 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
- 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, UK
- 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 Jap* 82:1599–1628
- Wetherald RT, Manabe S (1995) The mechanisms of summer dryness induced by greenhouse warming. *J Clim* 8:3096–3108