REPORT: Regional climate projections and bias correction methods over the Greater Alpine Region

M. Turco[1,*], J. von Hardenberg[1], E. Palazzi[1], and A. Provenzale[2]

[1]Institute of Atmospheric Sciences and Climate (ISAC), National Research Council (CNR), Torino, Italy.
[2]Institute of Geosciences and Earth Resources (IGG), National Research Council (CNR), Pisa, Italy.
[*]e.mail: m.turco@isac.cnr.it

March 30, 2015

1 Introduction

Global Circulation Models (GCMs) are tools of primary importance to obtain future climate projections under different anthropogenic forcing scenarios. However, the GCM coarse resolution (generally a few hundred kilometres) is not suitable for analysing the projections on a regional scale (generally a few tens of kilometres). Regional climate projections are necessary to assess the impacts of the potential future climate on ecosystems, hydrology and for performing risk assessments and, indeed, their development is a strategic topic in national and international climate programs (see, e.g. the WCRP CORDEX initiative *Giorgi et al.*, 2009). The gap between the coarse resolution GCMs and the scales at which most impacts happen is usually bridged by means of dynamical and/or statistical downscaling techniques (*Giorgi and Mearns*, 1991; *Wilby et al.*, 2004; *Benestad et al.*, 2008).

The dynamical downscaling approach is based on high-resolution Regional Climate Models (RCM), with typical resolutions of tens of kilometres, driven by GCMs within a limited domain (*Giorgi and Mearns*, 1991). The RCMs explicitly solve mesoscale atmospheric processes and provide spatially coherent and physically consistent output. However, they have in general considerable biases and therefore often cannot be directly used as input for impact models but must be corrected/calibrated. Several post-processing methods have been proposed to reduce the biases of the RCMs (see, e.g. the reviews of *Maraun et al.*, 2010 and of *Teutschbein and Seibert*, 2012). Although they may improve climate projections in most cases (*Maraun*, 2012), these techniques are critically sensitive to the assumption of the bias stationarity between the control and transient runs, which cannot be taken for granted (Christensen et al., 2008). This limitation should be taken into account for climate change studies (see *Ehret et al.*, 2012, and references therein for a debate on this topic).

In this work, we focus on precipitation and temperature, two key variables for many impact sectors (agriculture, hydrology, etc.) and we describe a relatively simple strategy to develop high-resolution projections over the Greater Alpine Region (GAR) combining dynamical downscaling and statistical methods following the Model Output Statistics approach (MOS; see *Maraun et al.*, 2010 for more details). We use the RegCM4 regional climate model run by the ICTP Group (see `http://www.medcordex.eu/` for more details). The data are available for the period 1970-2005 (historical run) and for the period 2006-2100 (RCP8.5 scenario) at the web page of the NextData project (`http://www.nextdataproject.it/`). For precipitation, which is typically highly variable in this region and has a large uncertainty in RCMs, we compare the outputs simulated by the RegCM4 model with those obtained applying three bias-correction methods, in order to disentangle the advantages and shortcomings of these approaches. For temperature, we implement and evaluate a simple bias-correction method for the maximum and minimum temperature simulated by the RegCM4 model. To test the robustness of the post-processing methods in a changing climate, we apply a cross-validation approach in which we split the period of the historical run into two subsets, one used to calibrate the bias-correction schemes, and the other to evaluate them.

The study is organized as follows: after this introduction, section 2 describes the data and methods used, section 3 analyses the validation results in the control period while section 4 analyses future projections. Finally, section 5 summarizes the main results of this study.

2 Data and methods

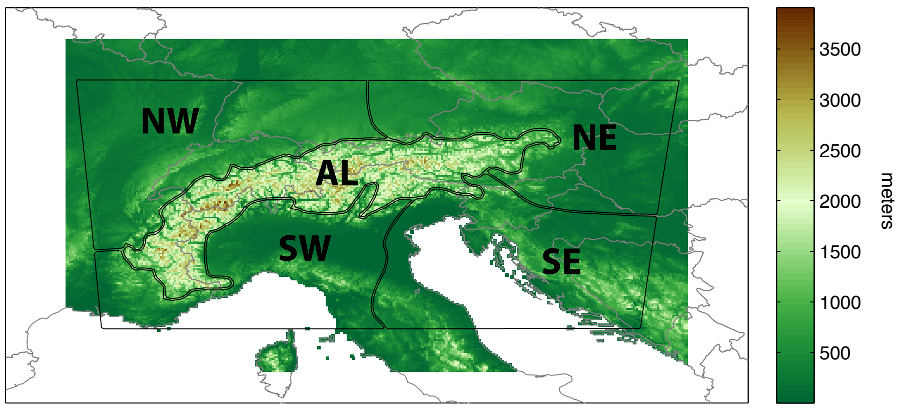

Figure 1: Orography of the domain of study and climate homogeneous sub-regions (as in *Haslinger et al.*, 2013).

We focus on the Greater Alpine Region (GAR; $2 - 17.5°$ E, $43 - 49°$ N; Figure 1). This region, which is geographically complex and heterogeneous, is characterized by a great climate variability due to the influences of different climatological regimes, such as the Atlantic, Mediterranean, continental, and polar regimes. In particular, the GAR region is characterized by large mountain chains, with the highest peaks reaching up to 4,000 m and deep valleys that act directly on the precipitation and temperature regimes. Due to this variability, this region is a challenging area for the climate models, since they must be able to simulate very different climates in a relatively small area with notable orographic complexity (Figure 1).

The Med-COordinated Regional climate Downscaling EXperiment (Med-CORDEX), is a coordinated contribution to CORDEX, a framework within the World Climate Research Program (WCRP). HyMeX and MedCLIVAR international programs support the med-CORDEX initiative. An ensemble of regional simulations at a 0.44° of resolution using state-of-the-art RCMs have been produced in the framework of Med-CORDEX. In this work we use the RegCM4 model run by the ICTP Group (`http://www.medcordex.eu/`). The data are available for the period 1970-2005 (historical run) and for the period 2006-2100 (RCP8.5 scenario) at the web page of the NextData project (`http://www.nextdataproject.it/`).

The observed precipitation data used to calibrate and evaluate the RCM are provided by the recently developed gridded dataset *EURO4M-APGD* (*Isotta et al.*, 2014) available at a resolution of about 5 km. To compare the simulations with the observations, we applied an up-scaling process to the EURO4M-APGD grid at 0.44° of resolution by means of an interpolation process. This consists in (i) bilinearly interpolating the original data to a 1 km scale, then (ii) aggregating these series to 0.44° of resolution averaging the 1 km points that fall in the respective pixel. The observed data are available for the period 1971-2008.

For temperature, we consider the high-resolution dataset of interpolated observations called EOBS version 10.0 (*Haylock et al.*, 2008). This dataset interpolates daily temperature data from quality-controlled stations over a regular grid of 0.44° resolution; data are available for the period 1950-2013 (see `http://eca.knmi.nl/` for more details).

For precipitation, in addition to evaluating the RCM model outputs, we evaluate three bias-correction procedures on a monthly scale: (1) $BIAS$, which consists in scaling the simulation by the ratio of the observed and simulated mean in the train period (*Teutschbein and Seibert*, 2012); (2) $QQMAP$, which calibrates the simulation removing quantile-dependent biases (*Wilcke et al.*, 2013); (3) $QQPAR$, which is based on the assumption that both the observed and simulated intensity distributions are well approximated by the gamma distribution, it is therefore a parametric q-q map that uses the theoretical instead of the empirical distribution (*Piani et al.*, 2010). For temperature, we evaluate the RCM model outputs and a simple bias-correction procedure, which consists in adding to the daily RCM simulation the mean difference between the observations and the simulation in the training period. To do so, we used the bias-correction implementation provided in the MLToolbox package (`https://meteo.unican.es/trac/MLToolbox/wiki`).

To check the robustness of the statistical methods to changing climate conditions, we determine the correction parameters on one subset of the data (the calibration period is 1971-1987 for precipitation and 1970-1987 for temperature), and validate the methods on the other (the test period is 1988-2005 for precipitation and for temperature). That is, we verify the ability of these post-processing methods to reduce the systematic error in the RCM variable from the knowledge of climatic data outside the period used to train the models.

Label	**Description**	**Units**
PRCPTOT	total precipitation	mm
SDII	Mean precipitation amount on a wet day	mm
CDD	consecutive dry days ($< 1mm$)	days
CWD	consecutive wet days ($> 1mm$)	days
R10	number of days with precipitation over 20 mm/day	days
R20	number of days with precipitation over 20 mm/day	days
RX5DAY	maximum precipitation in 1 day	mm
TXM	Mean of Maximum Daily Temperature	°C
TXX	Maximum of Maximum Daily Temperature	°C
SU	Summer Days - Number of Days with Tx>25°C	days
TNM	Mean of Minimum Daily Temperature	°C
TNN	Minimum of Minimum Daily Temperature	°C
TR	Tropical Nights - Number of Days with Tn>20°C	days

Table 1: Climatic mean and extreme ETCCDI indices for precipitation and for maximum and minimum temperature used in this work.

To evaluate the similarities between the simulated (both RCM outputs and bias-corrected ones) and observed climatologies, we consider standard verification measures (e.g. calculation of the bias) of a subset of the standard ETCCDI indices (Table 1), characterizing mean and extremes values (*WMO*, 2009). In general, all the indices are calculated on annual, seasonal and monthly scales over the test period 1986-2005 (except those which make no sense on a seasonal scale, such as the CDD index, since the longest period without rainfall can occur across two seasons). The comparisons between the simulated and observed seasonal climatologies are then resumed by the Taylor diagram (*Taylor*, 2001). This diagram makes possible a summary of three metrics of spatial similarity in a single bidimensional plot: standard deviation (S), centred root-mean-square difference (R) and correlation (C). Two variations from the standard Taylor diagram have been applied in this study:

- The statistics were normalized dividing both the centred root mean square error and the standard deviations of the simulated fields by the standard deviation of the observations. In this way it is possible to compare the different indices.

- To include information about overall biases (M), the colour of each point indicates the difference between the simulated and observed values.

3 Verification results

3.1 Precipitation

In Figures 2 we show the comparison maps for the simulated and the corresponding bias-corrected values (BIAS method) at annual scale for the precipitation indices of Table 1. The panels in this figure show the annual values of the indices (averaged during the test period 1986-2005) for the gridded observational dataset (first column), the RCM simulated values (second column), and the MOS values (third column). This figure shows that the bias-corrected values generally outperform the uncalibrated RCM outputs (except for the CDD index).

The Taylor diagrams of Figures 3, 4 and 5 summarize the verification results for all the seasons, indices and methods for the indices of Table 1. Squares indicate the RCM, while dots indicate the MOS values. The numbers indicate the different climate indices considered. Simulated results that agree well with observations will lie nearest the point marked "OBS" on the x-axis. This figure shows that, overall, all the methods considerably improve the RCM results for all the indices, with larger spatial correlation values and standard deviation closer to the observed ones. In terms of both bias (M) and centred root mean square error (R), the MOS methods generally improve the RCM results. Better results are achieved for the $QQMAP$ and $QQPAR$ methods.

In order to assess the correspondence of the simulated and observed annual cycles, we analyse the performance of the methods in the different sub-regions (according to Figure 1) on a monthly scale. Figure 6 shows two examples for the MOS correction based on the simple BIAS method (which consists in scaling the simulation by the ratio of the observed and simulated mean in the training period). These two cases illustrate the potential limitations of the bias-correction methods. A crucial assumption of all bias-correction methods is that RCM biases be constant over time. We assess the validity of this hypothesis by directly comparing the observed annual cycle with corresponding model outputs over two historical periods. The panels in first column of Figure 6 display the annual cycle, observed and simulated by the RCM, for the period 1971-1987. This is the period used to calibrate the bias correction. The second column shows the comparison for the period 1988-2005, which is the period used to test the bias correction. Finally, the third column shows the comparison between the observations and the values obtained after the bias correction. The three panels in the first row of Figure 6 refer to the verification results for the south-eastern sub-area of the GAR region (SE, see Figure 1). For this sub-region, the RCM is able to reproduce reasonably well the annual cycle, and its bias is approximately constant in the two periods except in January, February and March. Consequently, the MOS calibration based on the simple bias correction amplifies, instead of reducing, the biases during these months. The second row refers to another illustrative example (sub-area in the North-East of the GAR region, NE), in which the RCM presents a large overestimation during the winter/spring months with respect to the observed values, although these biases are broadly constant over the two periods. In this case, the bias correction is able to reduce these errors. That is, in the SE region the RCM shows relatively low biases but these are non-constant in time and consequently the MOS method leads to an increase in the monthly biases. On the other hand, in the NE region, the RCM shows larger biases that are broadly constant in time and consequently the MOS method is able to reduce them. Similar results have been found in the frequency of rainy days ($R1$) index, for other sub-areas and for the other two methods. Overall, these outcomes indicate that the uncertainty related to the assumption of bias stationarity is not negligible and that the regional projections of precipitation, both from direct model outputs or bias-correction values, must be used and interpreted very carefully.

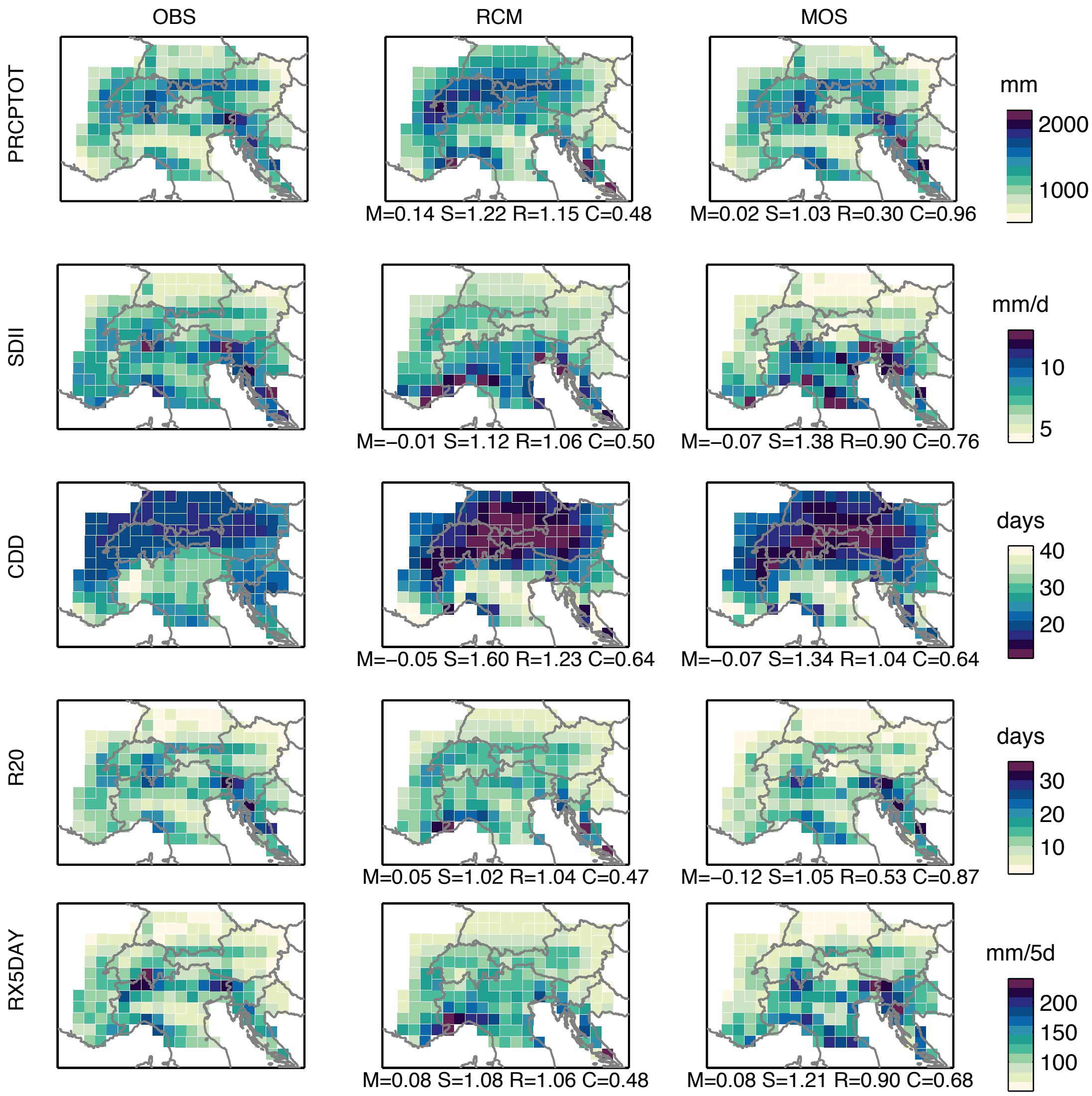

Figure 2: Spatial distribution of the (left) observed, (middle) RCM and (right) MOS (BIAS method) mean values (averaged over the test period: 1988-2005) for representative precipitation indices shown in Table 1. The spatial validation scores (correlation and errors) for the MOS and RCM simulated values are given below the corresponding panels: bias (or mean error M), relative standard deviations (S), correlation (C) and centred root-mean-square (R).

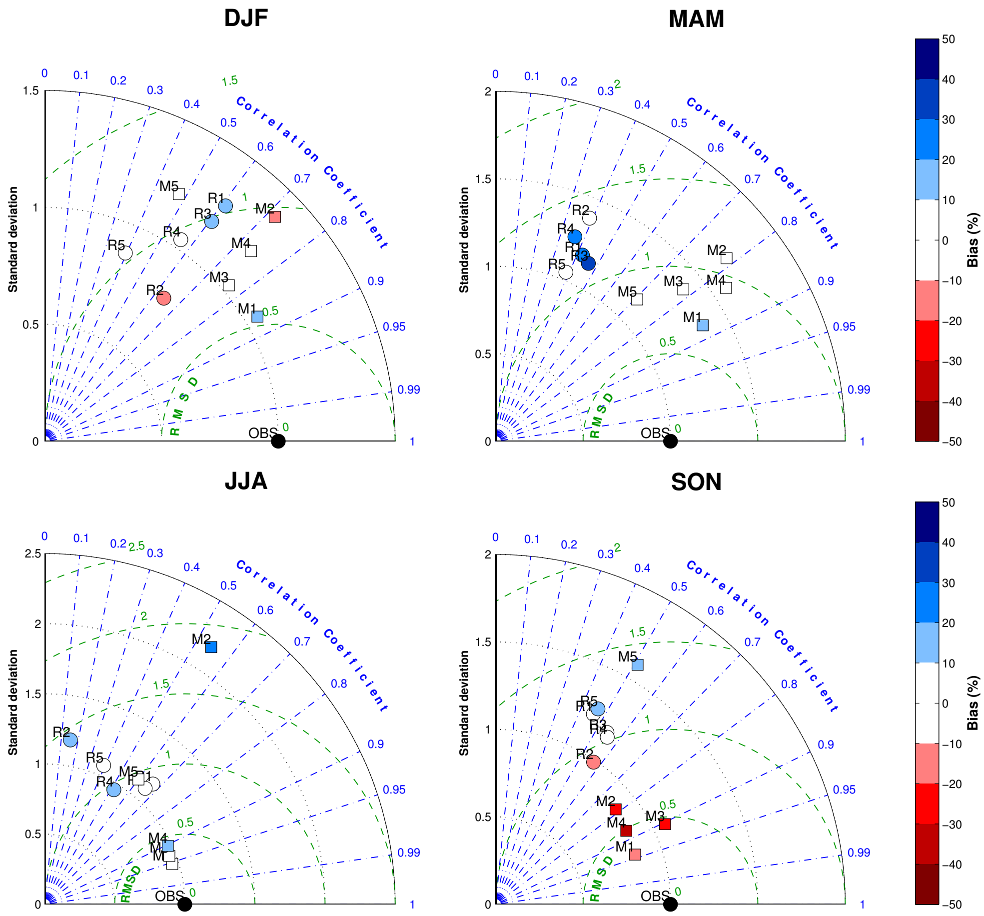

Figure 3: Taylor diagrams for the seasonal precipitation climatology of the indices shown in Table 1. The numbers correspond to the different indices. Better results are closer to observation (OBS). The circles are used for the uncalibrated RCM model, while the squares for the BIAS method. 1= PRCPTOT; 2= SDII; 3=R10; 4=R20; 5=RX5DAY. The colours indicate the bias (in percentage respect to the observed mean).

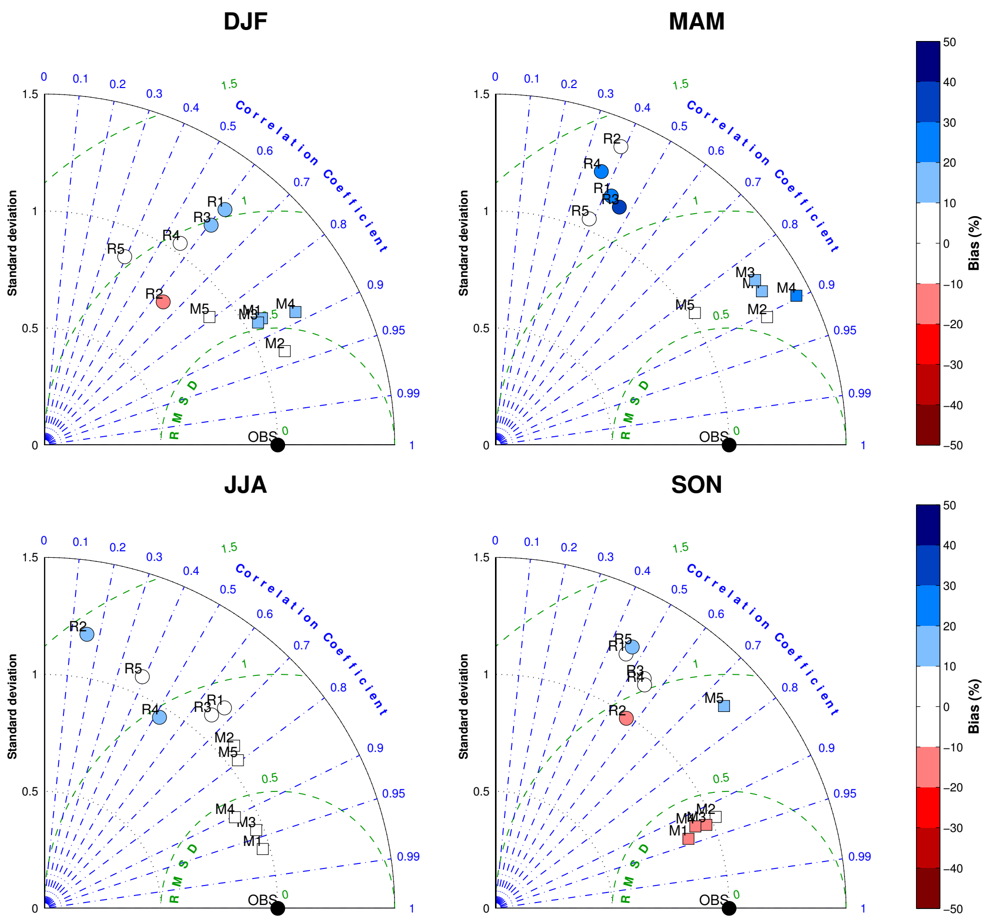

Figure 4: Same as Figure 9 but for QQMAP method.

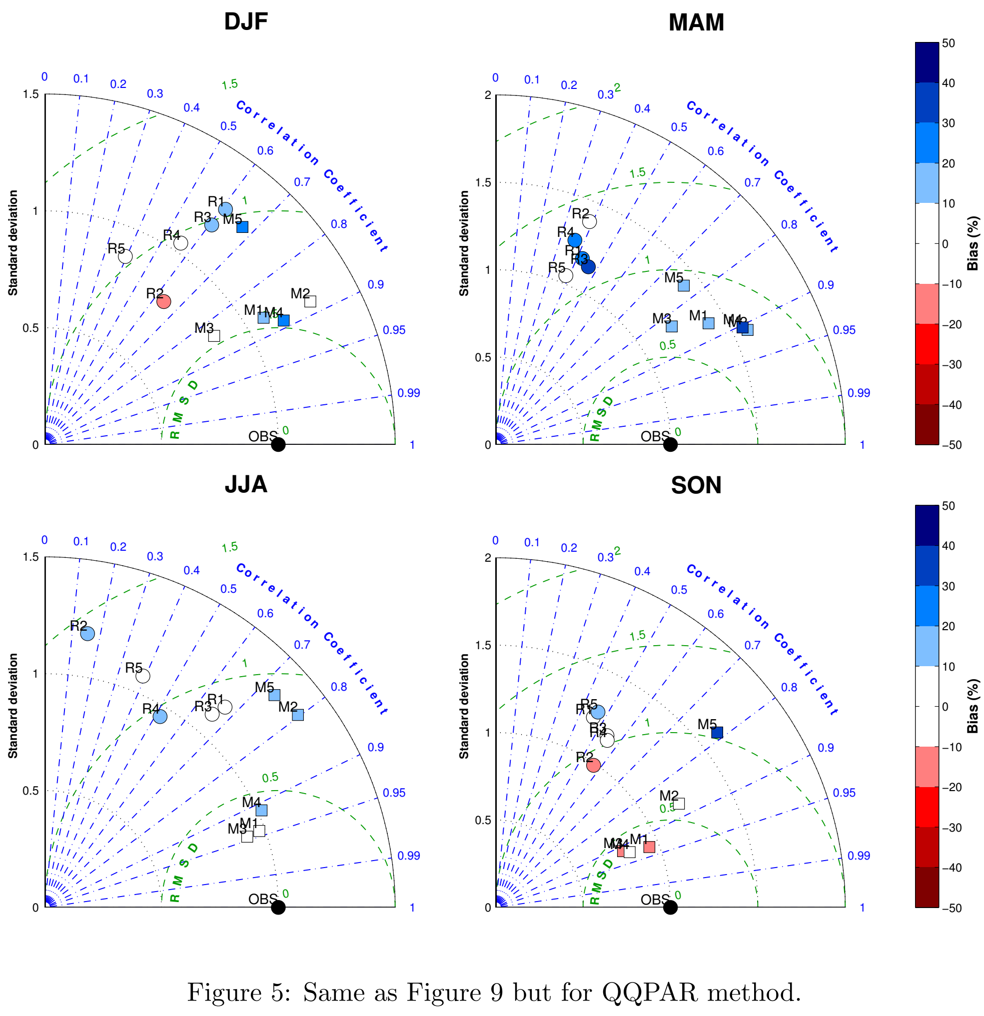

Figure 5: Same as Figure 9 but for QQPAR method.

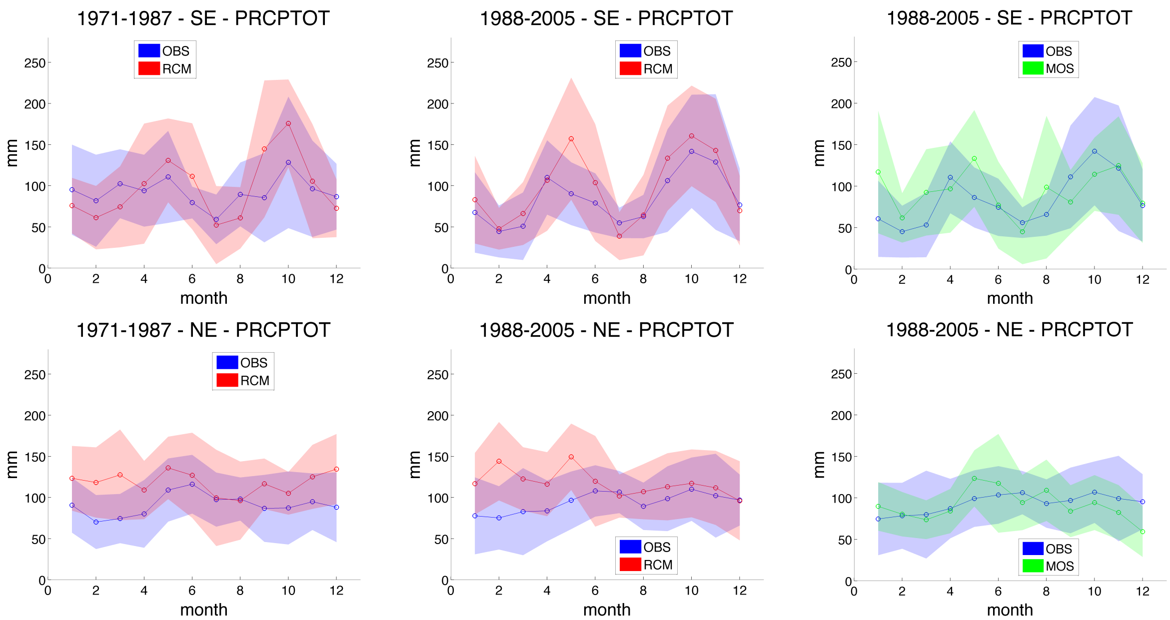

Figure 6: Seasonal cycle of the monthly precipitation for two representative sub-areas basin (according to Figure 1, South-East (first row; SE) and North-East (second row; NE) for the training period (1971-1987, left panels) and the test period (1988-2005, right panels). The blue shaded band spans the values for the observed data ($\pm$ 1 σ), the red one spans the RCM values while the green one spans the MOS values (BIAS method).

3.2 Temperature

Figure 7 shows the comparison maps for the simulated and the corresponding bias-corrected temperatures on an annual scale. The panels in this figure show the annual values of the indices (averaged during the test period 1988-2005) for the E-OBS dataset (first column), the RCM simulated values (second column), and the bias-corrected values (third column). This figure shows that the MOS values generally outperform the uncalibrated RCM outputs.

The Taylor diagrams in Figure 8 summarize the verification results for all the seasons and indices. This figure shows that, overall, the MOS method considerably improves the RCM results for all the indices, with larger spatial correlation values and standard deviation closer to the observed ones. The MOS method generally reduces underestimation of the dynamical model.

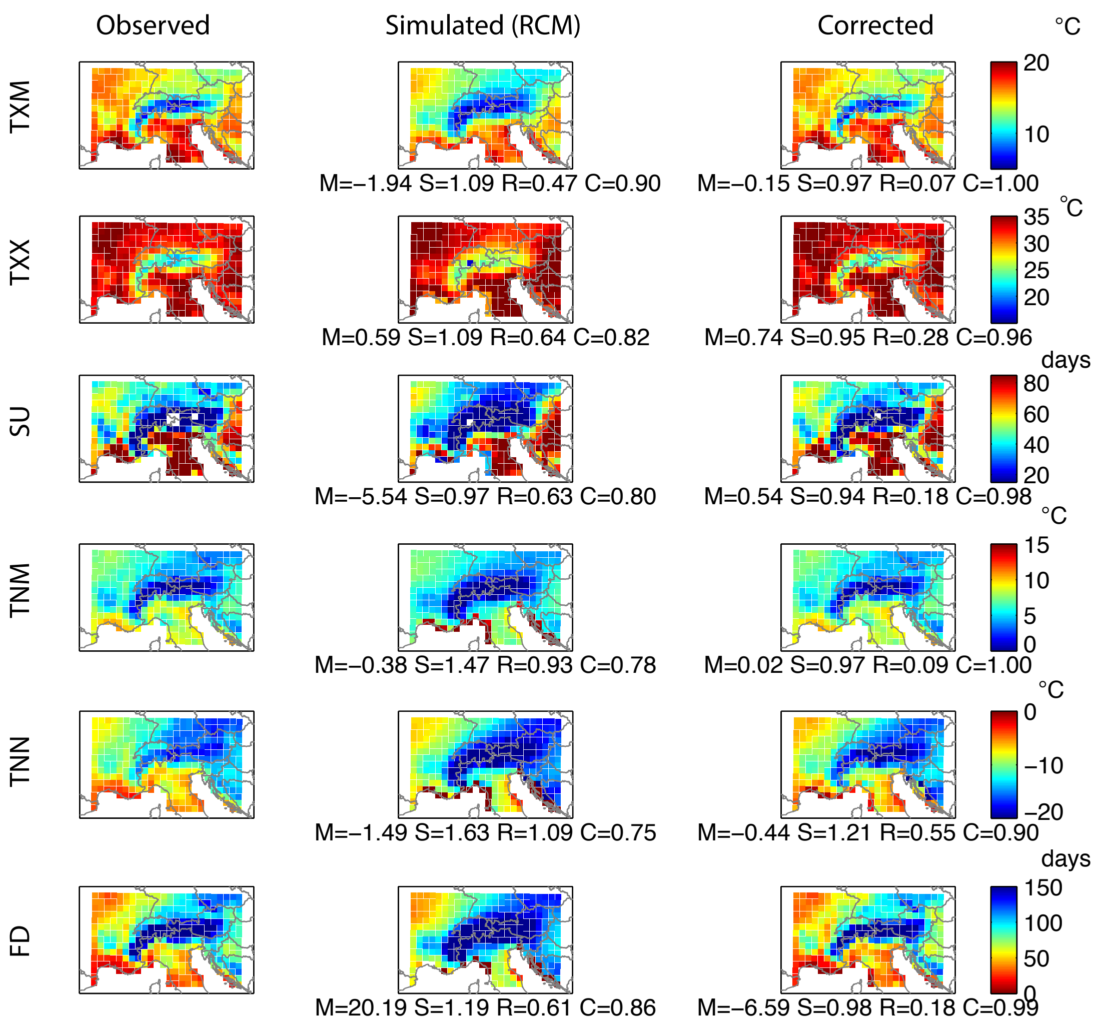

Figure 7: Spatial distribution of the (left) observed, (middle) RCM and (right) corrected (BIAS method) mean values (averaged over the validation period: 1988-2005) for representative temperature indices shown in Table 1. The spatial validation scores (correlation and errors) for the MOS and RCM simulated values are given below the corresponding panels: bias (or mean error M), relative standard deviations (S), correlation (C) and centred root-mean-square (R).

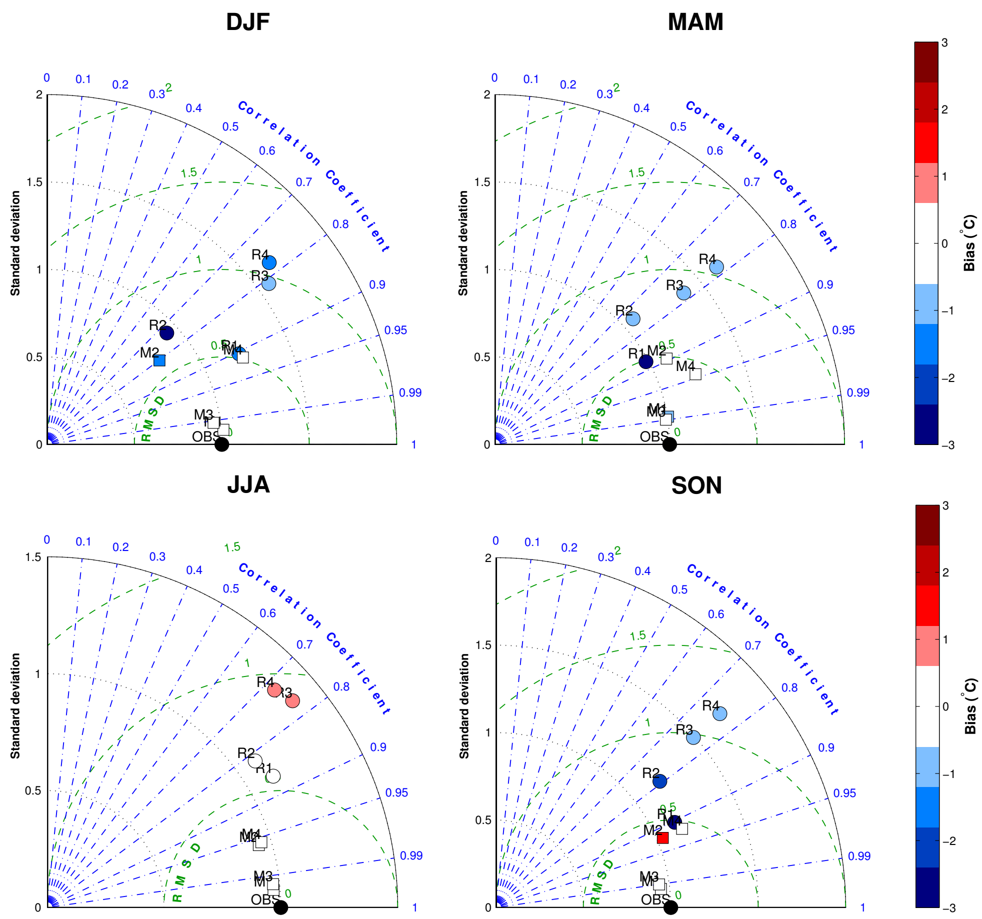

Figure 8: Taylor diagrams for the seasonal temperature climatology of the indices shown in Table 1. The numbers correspond to the different indices. Better results are closer to observation (OBS). The circles are used for the uncalibrated RCM model, while the squares for the BIAS method. 1= TXM; 2= TXX; 3=TNM; 4=TNN. The colours indicate the bias (i.e. the difference respect to the observed mean).

Finally, to assess the correspondence of the simulated and observed annual cycles, we analyse the performance of the post-processing method for temperature in the different sub-regions (according to Figure 1) on a monthly scale. Figure 9 shows the annual cycle of the TXM index. Generally, both the RCM and the MOS method reproduce well the annual cycle, with lower bias, as expected, for MOS. The correction is less effective over the Alps. Similar results are also found for the other indices (not shown). In contrast to the results for precipitation, temperature biases have been found to be relatively stationary, suggesting that the MOS approach for temperature is a reasonable solution to provide inputs for impact studies.

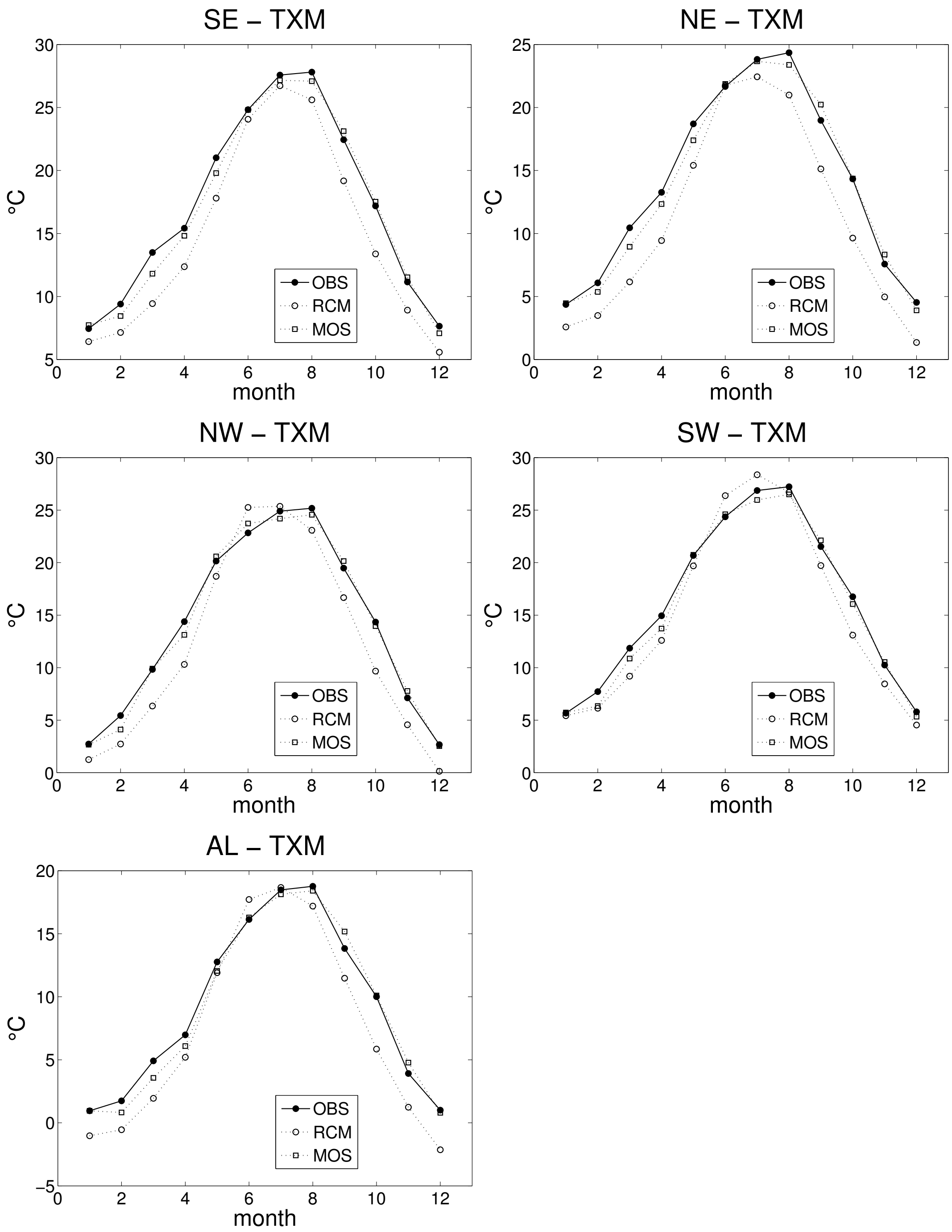

Figure 9: Seasonal cycle of the monthly mean of the maximum temperature for each sub-areas basin (according to Figure 1). The black line represents the observed climatology. The dotted line with circle indicates the values for the RCM while the square one shows the calibrated values by mean of a simple BIAS correction.

4 Future projections

In this section, the future projections obtained from the RCM and the MOS methods are presented. Here the focus is on analysing the impact of the MOS method on the climate change signals (CCS). CCS is calculated as the difference between future (time average over the period 2070-2099) and present (1971-2000 average) simulations.

Figures 10 and 11 show that the future projections for precipitation and temperature, respectively, are quite similar between the RCM and the MOS values. These results suggest that these post-processing techniques are able to preserve the climate change signal of the RCM. The analysis on a seasonal scale confirms that the climate change signal is preserved (not shown).

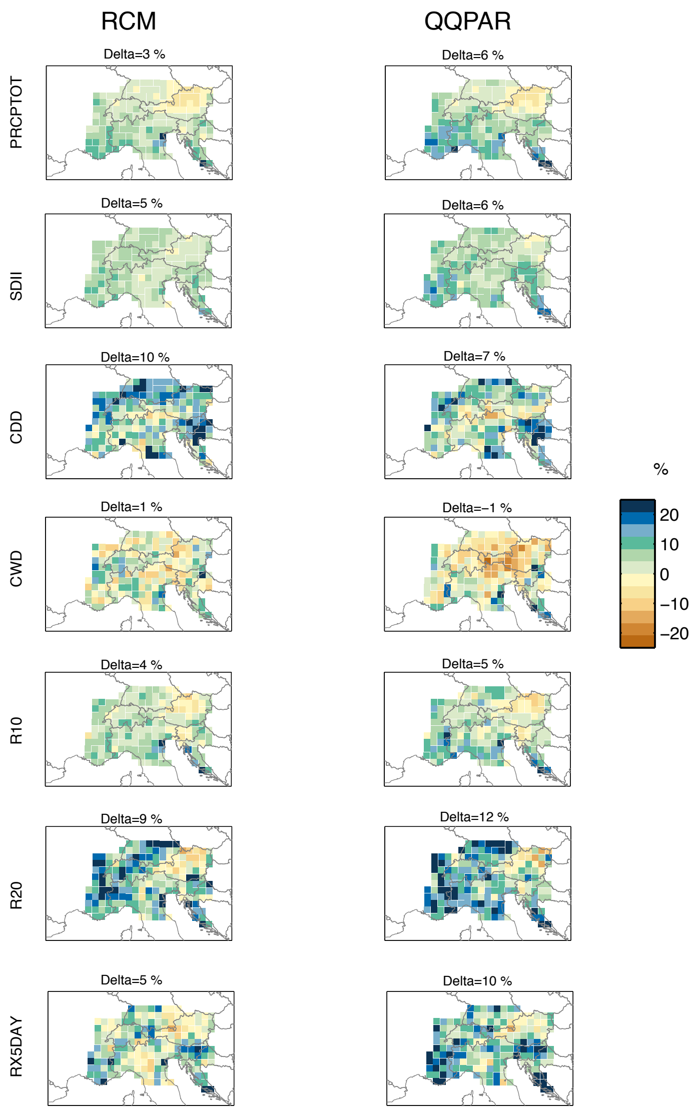

Figure 10: Climate change projections for (left panels) RCM and (right panels) corrected (BIAS method), given by the difference of future (2070-2099) and present (1971-2000) simulations, for the indices shown in Table 1.

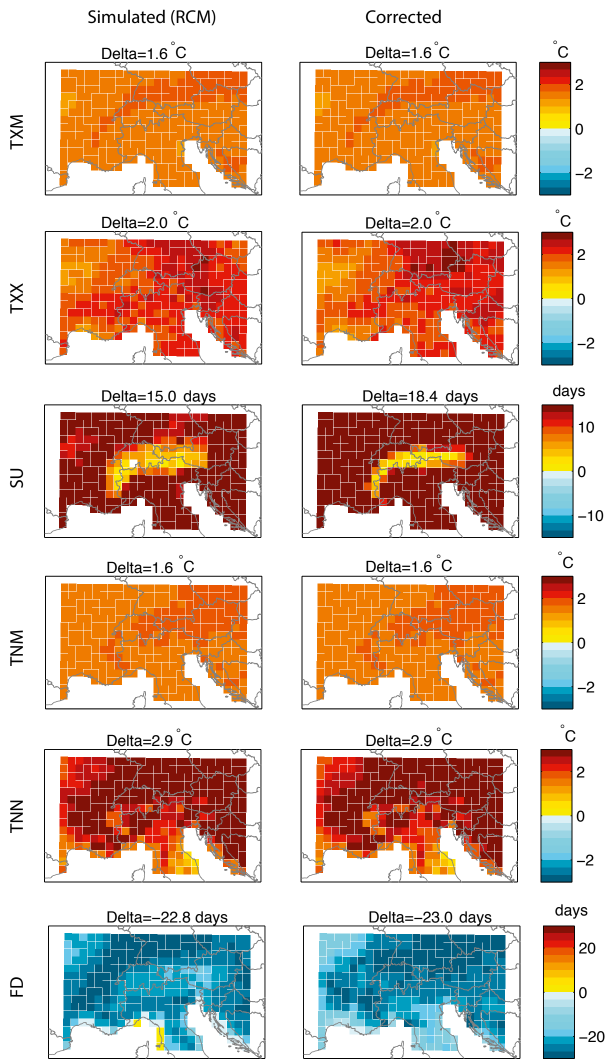

Figure 11: Climate change projections for (left panels) RCM and (right panels) corrected (BIAS method), given by the difference of future (2070-2099) and present (1970-1999) simulations, for the indices shown in Table 1.

5 Conclusions

In this work, a relatively simple strategy to develop high-resolution projections over the Greater Alpine Region (GAR) combining dynamical downscaling and statistical methods has been presented and evaluated. We have focused on precipitation and temperature, two key variables for many impact sectors. We have evaluated different bias-correction methods to refine the precipitation and the temperature simulated by the RegCM4 model, assessing the strengths and limitations of this approach when applied for climate change studies.

The RCM model, evaluated during the historical period, shows reasonable skill in reproducing the spatial climatological pattern of different precipitation and temperature indices. However, this model has systematic biases that may reduce its applicability for impact models. The bias-correction methods we implement are usually able to reduce the systematic errors of the model, especially considering temperature indices. The correction is less effective for precipitation, also because the precipitation biases are non-stationary. This limitation is a common weakness of all the downscaling methods. Thus we stress the importance of carefully considering this additional source of uncertainty across the modelling chain connecting large-scale models and impact models.

The bias corrected outputs (at 0.44°) are available at the web page of the NextData project (`http://www.nextdataproject.it/`) in the standard netcdf format. The downscaled generated data are described in the Table 2.

File name	**Variable**	**Method**	**Period**
pr_BIAS_MED-44_HadGEM2-ES_historical_r1i1p1_ICTP-RegCM4-3_v7_day_19710101-20051231.nc	precipitation	$BIAS$	historical
pr_BIAS_MED-44_HadGEM2-ES_rcp85_r1i1p1_ICTP-RegCM4-3_v7_day_20060101-20991231.nc	precipitation	$BIAS$	future
pr_QQMAP_MED-44_HadGEM2-ES_historical_r1i1p1_ICTP-RegCM4-3_v7_day_19710101-20051231.nc	precipitation	$QQMAP$	historical
pr_QQMAP_MED-44_HadGEM2-ES_rcp85_r1i1p1_ICTP-RegCM4-3_v7_day_20060101-20991231.nc	precipitation	$QQMAP$	future
pr_QQMAP_MED-44_HadGEM2-ES_historical_r1i1p1_ICTP-RegCM4-3_v7_day_19710101-20051231.nc	precipitation	$QQPAR$	historical
pr_QQMAP_MED-44_HadGEM2-ES_rcp85_r1i1p1_ICTP-RegCM4-3_v7_day_20060101-20991231.nc	precipitation	$QQPAR$	future
tasmax_BIAS_MED-44_HadGEM2-ES_historical_r1i1p1_ICTP-RegCM4-3_v7_day_19700101-20051231.nc	maximum temperature	$BIAS$	historical
tasmax_BIAS_MED-44_HadGEM2-ES_rcp85_r1i1p1_ICTP-RegCM4-3_v7_day_20060101-20991231.nc	maximum temperature	$BIAS$	future
tasmin_BIAS_MED-44_HadGEM2-ES_historical_r1i1p1_ICTP-RegCM4-3_v7_day_19700101-20051231.nc	minimum temperature	$BIAS$	historical
tasmin_BIAS_MED-44_HadGEM2-ES_rcp85_r1i1p1_ICTP-RegCM4-3_v7_day_20060101-20991231.nc	minimum temperature	$BIAS$	future

Table 2: Description of the bias corrected outputs.

References

Benestad, R., I. Hanssen-Bauer, and D. Chen (2008), *Empirical-Statistical Downscaling*, World Scientific Publishers.

Ehret, U., E. Zehe, V. Wulfmeyer, K. Warrach-Sagi, and J. Liebert (2012), Hess opinions "should we apply bias correction to global and regional climate model data?", *Hydrology and Earth System Sciences*, *16*(9), 3391–3404.

Giorgi, F., and L. Mearns (1991), Approaches to the Simulation of Regional Climate Change - A review, *Reviews of Geophysics*, *29*(2), 191–216.

Giorgi, F., C. Jones, and G. Asrar (2009), Addressing climate information needs at the regional level: the CORDEX framework, *WMO Bulletin*, *58*(3), 175–183.

Haslinger, K., I. Anders, and M. Hofsttter (2013), Regional climate modelling over complex terrain: an evaluation study of COSMO-CLM hindcast model runs for the Greater Alpine Region, *Climate Dynamics*, *40*(1-2), 511–529.

Haylock, M. R., N. Hofstra, A. M. G. Klein Tank, E. J. Klok, P. D. Jones, and M. New (2008), A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006, *Journal of Geophysical Research*, *113*(D20), 1–12.

Isotta, F. A., C. Frei, V. Weilguni, M. Perec Tadi, P. Lassgues, B. Rudolf, V. Pavan, C. Cacciamani, G. Antolini, S. M. Ratto, M. Munari, S. Micheletti, V. Bonati, C. Lussana, C. Ronchi, E. Panettieri, G. Marigo, and G. Vertanik (2014), The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data, *International Journal of Climatology*, *34*(5), 1657–1675.

Maraun, D. (2012), Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation sums, *Geophysical Research Letters*, *39*(6), 1–5.

Maraun, D., F. Wetterhall, A. M. Ireson, R. E. Chandler, E. J. Kendon, M. Widmann, S. Brienen, H. W. Rust, T. Sauter, M. Themessl, V. K. C. Venema, K. P. Chun, C. M. Goodess, R. G. Jones, C. Onof, M. Vrac, and I. Thiele-Eich (2010), Precipitation Downscaling under Climate Change: Recent Developments to Bridge the Gap Between Dynamical Downscaling Models and the End User, *Reviews of Geophysics*, *48*, 1–34.

Piani, C., J. Haerter, and E. Coppola (2010), Statistical bias correction for daily precipitation in regional climate models over Europe, *Theoretical and Applied Climatology*, *99*(1-2), 187–192.

Taylor, K. E. (2001), Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, *106*(D7), 7183–7192.

Teutschbein, C., and J. Seibert (2012), Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, *Journal of Hydrology*, *456457*(0), 12 – 29.

Wilby, R., S. Charles, E. Zorita, and B. Timbal (2004), Guidelines for use of climate scenarios developed from statistical downscaling methods, *Tech. rep.*, IPCC.

Wilcke, R. A. I., T. Mendlik, and A. Gobiet (2013), Multi-variable error correction of regional climate models, *Climatic Change*, *120*(4), 871–887.

WMO (2009), Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation, *Tech. Rep. WCDMP No. 72 WMO/TD-No. 1500*, WMO.