Hydrology and
Earth System
Sciences

Open Access

Regional climate models downscaling in the Alpine area with multimodel superensemble

D. Cane[1], S. Barbarino[1], L. A. Renier[1,2], and C. Ronchi[1]

[1]Regional Agency for Environmental Protection – Arpa Piemonte, Torino, Italy
[2]IPLA Spa, Torino, Italy

Correspondence to: D. Cane (daniele.cane@arpa.piemonte.it)

Abstract. The climatic scenarios show a strong signal of warming in the Alpine area already for the mid-XXI century. The climate simulations, however, even when obtained with regional climate models (RCMs), are affected by strong errors when compared with observations, due both to their difficulties in representing the complex orography of the Alps and to limitations in their physical parametrization.

Therefore, the aim of this work is to reduce these model biases by using a specific post processing statistic technique, in order to obtain a more suitable projection of climate change scenarios in the Alpine area.

For our purposes we used a selection of regional climate models (RCMs) runs which were developed in the framework of the ENSEMBLES project. They were carefully chosen with the aim to maximise the variety of leading global climate models and of the RCMs themselves, calculated on the SRES scenario A1B. The reference observations for the greater Alpine area were extracted from the European dataset E-OBS (produced by the ENSEMBLES project), which have an available resolution of 25 km. For the study area of Piedmont daily temperature and precipitation observations (covering the period from 1957 to the present) were carefully gridded on a 14 km grid over Piedmont region through the use of an optimal interpolation technique.

Hence, we applied the multimodel superensemble technique to temperature fields, reducing the high biases of RCMs temperature field compared to observations in the control period.

We also proposed the application of a brand new probabilistic multimodel superensemble dressing technique, already applied to weather forecast models successfully, to RCMS: the aim was to estimate precipitation fields, with careful description of precipitation probability density functions conditioned to the model outputs. This technique allowed for reducing the strong precipitation overestimation, arising from the use of RCMs, over the Alpine chain and to reproduce well the monthly behaviour of precipitation in the control period.

1 Introduction

The Alps are a region very sensitive to the impacts of climate change: the temperatures increased here more than the world average (Ciccarelli et al., 2008) and the projections for the XXI century show again an increase higher than the average.

Piedmont region is located in north-western Italy, at the south-western edge of the Alpine chain. The Environmental Protection Agency of Piedmont region is in charge of producing reliable scenarios of the variation of climatic parameters in the changing climate, to allow the evaluation of the impacts on mountain hydrology (project AC-QWA, www.acqwa.ch), on the wildfire potential (project ALPFFIRS, www.alpffirs.eu, for more details see Cane et al., 2013a), on the permafrost (project PERMANET www.permanet-alpinespace.eu/), on the Alpine lakes (project SILMAS, www.silmas.eu), on mountain biodiversity (EU-INTERREG project "Biodiversità una risorsa da conservare"), on heat waves in the Po Valley towns (Nicolella and Cane, 2012) and on any other study of impacts affecting the regional environment.

The temporal target of our work is the mid-XXI century, to drive conclusions that can be used for tangible adaptation measures to climate change in a reasonable time.

Then we chose to focus on a single scenario instead of a range of different scenarios as, for our time interval of interest, the largest variations occur among the different models, while the different scenarios do not differ so much (Randall et al., 2007).

In this work we applied multimodel techniques on several regional climate model (RCM) outputs, which are combined together to obtain collective evaluations.

Multimodel combination is a pragmatic approach to estimating model uncertainties and to make climate projections more reliable. Their use in the climatic simulations is recommended by the Intergovernmental Panel on Climate Change (Knutti et al., 2010). The simplest multimodel technique is the "Poor Man Ensemble", which is an average of different models, without any bias correction or weighting ("equal weighting"), while more sophisticated approaches suggest applying model weights according to some measure of performance ("optimum weighting"). The results confirm that equally weighted multimodels on average outperform the single models (Krishnamurti et al., 1999, 2000; Yun et al., 2003), and that projection errors can in principle be further reduced by optimum weighting. However, this not only requires accurate knowledge of the single model skill, but the relative contributions of the joint model error and unpredictable noise also need to be known to avoid biased weights (Weigel et al., 2010). Many weighting procedures were proposed on seasonal, decadal and climatic models (Giorgi and Mearns, 2002; Palmer et al., 2004; Coppola et al., 2010; Feng et al., 2011). Christensen et al. (2010) showed that the use of model weights is sensitive to the aggregation procedure and showed different sensitivities to the selected metrics. They did not find compelling evidence of an improved description of mean climate states using performance-based weights in comparison to the use of equal weights. They suggested that model weighting adds another level of uncertainty to the generation of ensemble-based climate projections, which should be suitably explored, although their results indicate that this uncertainty remains relatively small for the weighting procedures examined.

Our evaluation is based on regional climate models (RCMs) calculated by the EU project ENSEMBLES on the SRES scenario A1B basis: all the model runs refer to the same grid including Europe.

We established two study areas: the greater Alpine area (GAR, Fig. 1), including all the Alps (coordinates: 3.00–20.25° E, 41.50–51.25° N, and a smaller box covering Piedmont region (OI, Fig. 2) with higher resolution data (coordinates: 6.5625–9.4375° E, 44.0625–46.4375° N).

The reference observations for the GAR are extracted from the European gridded dataset E-OBS produced by the EU project ENSEMBLES (Haylock et al., 2008). Daily temperature and precipitation observations (covering the period from 1961 to the present) which are derived from this dataset are available at a resolution of 25 km.
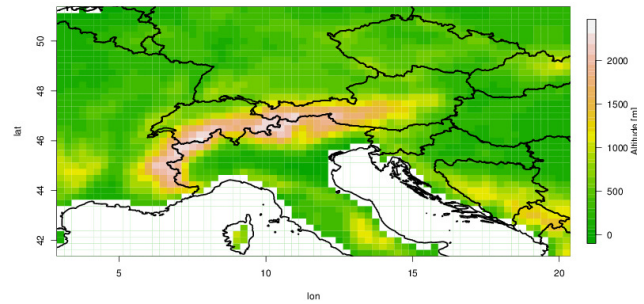
Fig. 1. The greater Alpine area map.

For the study area of Piedmont, daily temperature and precipitation observations (covering the period from 1957 to the present), collected by the Environmental Protection Agency of Piedmont region, were carefully gridded on a 14 km grid over Piedmont region by using an optimal interpolation (OI) technique. More details can be found in Sect. 2.

In the multimodel superensemble technique (Krishnamurti et al., 1999) the models are unbiased and weighted with an adequate set of weights calculated during the so-called training period, with comparison to the observations. This technique is widely applied to weather forecast models (an example in Piedmont can be found in Cane and Milelli, 2006) and to seasonal climate forecasts (Krishnamurti et al., 2000). The standard multimodel superensemble technique was here applied to the temperature fields regarding the period 1961–1980 as training dataset to calculate weights and to obtain daily fields of reanalyses (1981–2000) and scenarios (1981–2050).

A new probabilistic multimodel superensemble dressing, with careful description of precipitation probability density functions conditioned to the model outputs was applied to the precipitation fields. This technique allows for a better correction of precipitation biases depending on the value of the forecast precipitation. For more details, please see Sect. 2.

The multimodel uncertainty evaluation is usually addressed by multimodel uncertainty post-processors like the Bayesian Model Averaging (BMA) introduced by Raftery et al. (2003, 2005) and the Model Conditional Processor (MCP) developed by Todini (2008).

In this work, we applied a simple evaluation of multimodel uncertainty on the multimodel superensemble technique in the hypothesis of multi-normal distribution. In the case of the probabilistic multimodel superensemble dressing the model uncertainty was evaluated directly by using a MonteCarlo technique (see Sect. 2 for more details).

In Sect. 3 we described our results, first of all through a validation of the multimodel techniques in the control period in Piedmont region, and then evaluating the future scenario both in Piedmont and in the whole Alpine area.
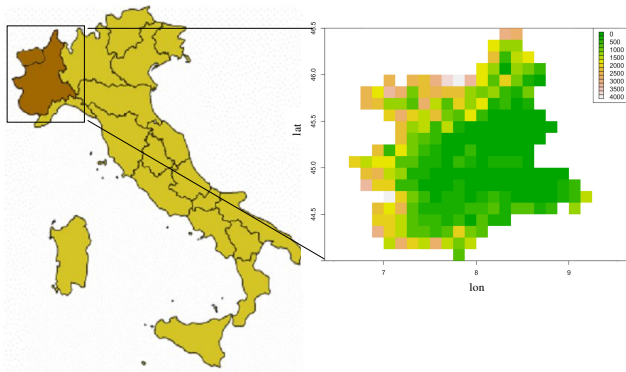
Fig. 2. The optimal interpolation map. In the detailed map the grid-point elevation (*m*) is showed.

2 Available data and description of the technique

2.1 Regional climate models and large scale observed data

The RCMs simulations used in this paper are a selection of 7 RCMs runs resulting from the ENSEMBLES project (Table 1), carefully chosen in order to maximise the variety of leading global climate models and of the RCMs themselves, and with a data amount compatible with our elaboration and storage facilities. Models descriptions can be found at http://ensemblesrt3.dmi.dk/.

For each model, the reanalysis runs from the ECMWF ERA-40 reanalysis (1961–2000) and the scenario runs (1961–2100) on SRES scenario A1B are available on a common grid at a resolution of 25 km.

We interpolated the daily data from the models on the GAR and OI domains as defined in the Introduction, with a simple bi-linear interpolation. The use of such an interpolating technique can introduce biases, but the multimodel techniques include a bias removal before applying the model average.

As for observation data, in this study we employed the E-OBS dataset, produced by the EU project ENSEMBLES. This dataset is an European land-only daily high-resolution gridded data set for precipitation and minimum, maximum, and mean surface temperature which spans from 1950 to the present. The data set has been designed to provide the best estimate of grid box averages rather than point values to enable direct comparison with RCMs (here the 25 km resolution dataset is employed). The authors employ a three-step process of interpolation, by first interpolating the monthly precipitation totals and monthly mean temperature using three-dimensional thin-plate splines, then interpolating the daily anomalies using indicator and universal kriging for precipitation and kriging with an external drift for temperature, then combining the monthly and daily estimates. Interpolation uncertainty is quantified by the provision of daily standard errors for every grid square.

2.2 Optimal interpolation of Piedmont data

Regarding the gridded dataset of daily temperature and precipitation data over Piedmont, an optimal interpolation (OI) technique was used to assimilate the row and high density ground station data, arbitrarily displaced in the region, on a selected regular three-dimensional grid map based on a background field (BF) (Kalnay, 2003).

Only for temperature, the background field is obtained by a linear tri-dimensional downscaling of ERA-40 archive spanning from 1957 to 2001 and of the ECMWF objective analysis from 2002 to 2009 on a selected grid (0.125° resolution, with careful description of the complex orography of the region).

The use of ERA-40 on the regional area is suggested by checking that the main climatological signals (trends, etc.) were congruent with the signals resulted from a station subset working in the period 1950–2000 in Piedmont (Ciccarelli et al., 2008). Where this preliminary congruence checking was not clear (i.e. for precipitation) the row station data themselves provided the background field at first level of gridding process.

The method enables to weight the contribute to the temperature/precipitation value on each grid point from the nearest observation data, through suitable parameters. A careful modulation of these parameters as a function of the data density and the use of an external background field help to achieve the time homogeneity and the spatial coherence of the final dataset.

2.3 Standard multimodel superensemble technique on temperature fields

The conventional superensemble forecast (Krishnamurti et al., 2000) constructed with bias-corrected data is given by

$$S = \overline{O} + \sum_{i=1}^{N} a_i \left(F_i - \overline{F_i} \right). \tag{1}$$

Where F_i is the ith model forecast, $\overline{F_i}$ is the mean of the ith forecast over the training period, $\overline{O}$ is the observed mean over the training period, a_i are regression coefficients obtained by a minimisation procedure during the training period, and N is the number of forecast models involved.

The calculation of the parameters a_i is given by the minimisation of the mean square deviation in the training period T. The sets of parameter a_i are estimated for each single pixel and not as an overall weighting of models as in BMA, hence the number of parameters is very high: the number of models times the number of pixels.

$$G = \sum_{k=1}^{T} (S_k - O_k)^2 \tag{2}$$

Table 1. The models used in the multimodel superensemble evaluation.

Acronym	Reg. Clim. Model	Global Clim. Model	Run by
DMI	HIRHAM5	Arpege	Danish Meteorological Institute
ICTP	REGCM3	ECHAM5	The Abdus Salam Intl. Centre for Theoretical Physics
HC	HadRM3Q0	HadCM3Q0	Hadley Centre for Climate Prediction and Research
CNRM	RM4.5	Arpege	Météo-France CNRM/GMGEC/EAC
ETHZ	CLM	HadCM3Q0	Swiss Institute of Technology (ETHZ)
KNMI	RACMO2	ECHAM5	The Royal Netherlands Meteorological Institute
MPI	REMO	ECHAM5	Max Plank Institute – Hamburg

By derivation $\left(\frac{\partial G}{\partial a_i} = 0\right)$, we obtain a set of N equations, where N is the number of models involved.

We then solve these equations using Gauss–Jordan method (Press et al., 1992).

A scheme of the technique can be found in Fig. 3.

The standard multimodel superensemble technique was applied to the temperature fields by using the period 1961–1980 as training dataset to calculate weights and to obtain daily fields of reanalyses (1981–2000) and scenarios (1981–2050).

ERA-40-driven RCMs have specific weights calculated from the comparison between their daily values in the training period and observations, GCM-driven RCMs have different weights calculated from their own values. The weights and biases are calculated independently for each model and grid-point, but do not depend on time (i.e. we applied the same bias correction and weights for any day of the year).

While in the ERA-40-driven model data we can expect a certain degree of correlation between model data and observations, in the scenario-driven models the difference between the model and the observed values are partly due to the specific parametrization of each numerical model and partly due to the necessary uncorrelation between GCM scenarios and reality. So, the weights we calculated are not employed to disentangle these contributions, but rather to evaluate which is the best way to average the two different contributions in order to obtain the best approximation of observations (always on average). While applying multimodel superensemble, we are making two strong hypotheses:

1. for any given point, the (annual averaged) biases of the models in the future scenario will remain the same way they were in the past scenario (training period);

2. for any given point, the weights of the models, hence their relative contribution to the final result, will remain the same. From our experience, if you are using fixed weights for a multimodel, it is better to have the largest possible training period; whereas if you are interested in a better correction, you must use short statistics close to the forecast period you are interested in (Cane and Milelli, 2006).
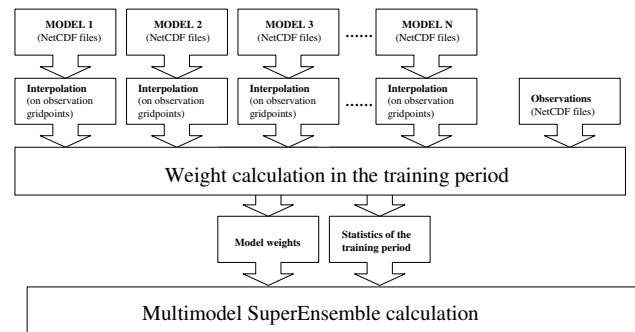
Fig. 3. Scheme of the standard multimodel superensemble technique.

We tested the temperature conditional probability density in the very begin of our work and, as expected, it is very close to a normal distribution. The combination of many models to a resulting multi-normal distribution can introduce heteroscedasticity in our multimodel: heteroscedasticity does not cause ordinary least squares coefficient estimates to be biased, although it can cause ordinary least squares estimates of the variance of the coefficients to be biased, possibly above or below the true population variance.

We evaluated the multimodel superensemble dressing uncertainty with a explicit calculation of the model variances and covariances in the training period, and combining them with the multimodel weights to obtain the final multimodel variance (here calculated in the hypothesis of non-independent models):

$$\sigma_{\mathrm{MM}}^2 = \sum_{i,j}^N a_i a_j \sigma^2 \left(F_i, F_j\right). \tag{3}$$

The use of any kind of metrics on multimodel evaluation highlights some features of the ensemble and carries with it some limitations (Knutti et al., 2010). Multimodel superensemble technique on RCMs data imposes a very simple metrics (a weighted average of un-biased models), which is quite straightforward and the results can be interpreted very easily, but brings some limitations about the use of the outputs:

1. the results should be bias-free on a long time scale and
   can be used averaged over long periods;

2. the extremes of the distribution can be underestimated
   or overestimated, hence the multimodel data cannot be
   used for the evaluation of distribution extremes;

3. the data are evaluated independently for each gridpoint.
   This means that nearby points can have un-correlated
   corrections, but at the same time the contribution of the
   data to the models via the multimodel permits to correct
   the scenarios in a very punctual position, hence allow-
   ing for a better representation of local features such the
   topography of the Alpine chain. The bias reduction ac-
   counts for the major part of the error reduction, while
   the better representation of the monthly cycle is mainly
   due to the averaging effect of the different models. In
   both cases, anyway, the correction is spread over a long
   time.

An evaluation of the models in the control period is available
in Sect. 3.

2.4 Probabilistic multimodel superensemble dressing on precipitation field

A new probabilistic multimodel superensemble dressing,
with ad hoc description of precipitation probability density
functions (PDFs) conditioned to the model outputs was ap-
plied to the precipitation fields.

We followed this algorithm:

1. We took the ERA40-driven RCMs and compared them
   with the observed precipitation on all grid points.

2. We considered all the days and points where the model
   produces a given precipitation.

3. We built the distribution of the observed values of that
   days/points (with bins of width ± 0.5 mm around the
   central value).

4. We repeated for any reasonable forecast value (up to
   300 mm day^{-1}).

5. We fitted the distributions so obtained with a set of func-
   tions, finding that the Weibull function is the best one to
   represent all of the distributions among a large set of
   possible candidates.

6. We interpolated and extrapolated the observed distribu-
   tions to obtain all possible distributions for all forecast
   values and we obtain the function PDF(F), where F is
   the precipitation forecast value.

7. We calculated the individual RCMs continuous ranked
   probability scores (CRPS) from ERA40-driven models.
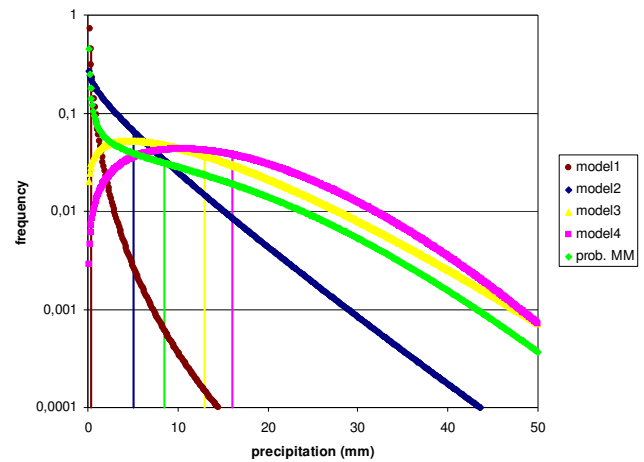
Fig. 4. Scheme of the probabilistic multimodel superensemble
dressing technique, example for a given day: probability density
function of observed data conditioned to different models and fore-
cast values, together with the final probability density function eval-
uated with multimodel.

8. We calculated the weights as the inverse of the CRPSs,
   and we normalized them according to Eq. 4, where a_i
   are now the weights of the multimodel superensemble
   dressing:

$$\frac{1}{a_i} = \frac{\frac{1}{\mathrm{CRPS}_i}}{\sum_i^N \frac{1}{\mathrm{CRPS}_i}}. \tag{4}$$

9. We applied the model-specific PDFs and the weights to
   the GCM-driven models to obtain weighted PDFs:

$$\mathrm{PDF_{MM}} = \sum_i^N a_i \mathrm{PDF}_i(F_i). \tag{5}$$

Figure 4 shows an example of our technique evaluated on
an ensemble of four models. For any given value of pre-
cipitation forecasted by the model, a model-specific PDF
is evaluated (here shown in brown, blue, yellow and pink),
and the final multimodel PDF is obtained with the correct
weights (in green). The vertical lines represent the original
"deterministic" value, while the green vertical line associ-
ated with the multimodel distribution is the average of the
multimodel PDF. The obtained average value can differ sig-
nificantly from the rude average of the input models (the so-
called Poor Man Ensemble), but the availability of a bias-
corrected PDF also allows for widening the ensemble spread,
trying to correct the under-dispersion of the multimodel en-
semble (in this example).

This technique allows for a better correction of precipita-
tion biases depending on the value of the forecast precipi-
tation. For more details and a verification on weather fore-
cast models, please refer to Cane and Milelli (2010). Multi-
model superensemble dressing precipitation data were used

successfully to drive an hydrological model and performed better than those obtained with the BMA approach (Cane et al., 2013b).

The probabilistic multimodel superensemble dressing technique was applied to the precipitation fields considering the period 1961–1980 as training dataset to calculate weights and to obtain daily fields of reanalyses (1981–2000) and scenarios (1981–2050).

For any given day a value is extracted randomly from the PDF to give a unique time series of precipitation. The use of a random extraction is justified by the large number of the samplings (∼ 25 000 in the considered period) and by the uncorrelation between the scenarios and the observations.

We addressed multimodel uncertainty using a MonteCarlo technique: we simulated 100 realizations of the multimodel precipitation from our empiric PDFs and we evaluated the statistics (average and quantiles) of monthly precipitation and yearly wet/dry periods (defined as the average number of periods with precipitation above/below 1 mm lasting more than 5 days).

In the case of precipitation, the model metrics is evaluated on ERA40-driven models and not on GCM-driven models: this is necessary, because we need to keep the correlation between forecast and observed precipitation but, at the same time, we can use these PDFs to quantify how much any single model has to be corrected in order to obtain more realistic precipitation values. On the other hand, CRPSs have been used to quantify how much any single model is able to reproduce the statistical distribution of the data. In this case we are confident to apply this two metrics evaluated on reanalyses to scenarios, because the under/overestimation of precipitation (described by PDF) will be the same for a given atmospheric pattern, both evaluated by ERA40 and by the GCM, while the CRPS is the ability of a given model to reproduce distributions and we think that we can use it to assess which model has to be ascribed the highest importance. In other words, once we estimated the bias correction of a given RCM via the calculation of its PDF on a realistic synoptic evaluation as represented by ERA40, we are assuming that the GCMs are able to reproduce the correct synoptic climatology, on the average, over a period of 20 yr: so we apply the weights calculated by ERA40 on the GCMs. Please notice that, while the weights of the standard multimodel superensemble techniques require the contemporary evaluation of any given model weight, in this case the CRPSs are evaluated independently, and only eventually we calculate the inverse and normalise these inverses to obtain the weighting.

Again, here we state the limitations of our technique:

1. the data must be evaluated on a monthly basis, with aggregation on a long period;

2. they cannot be used for extreme precipitation events because of the averaging effect (although reduced by the use of the PDFs);
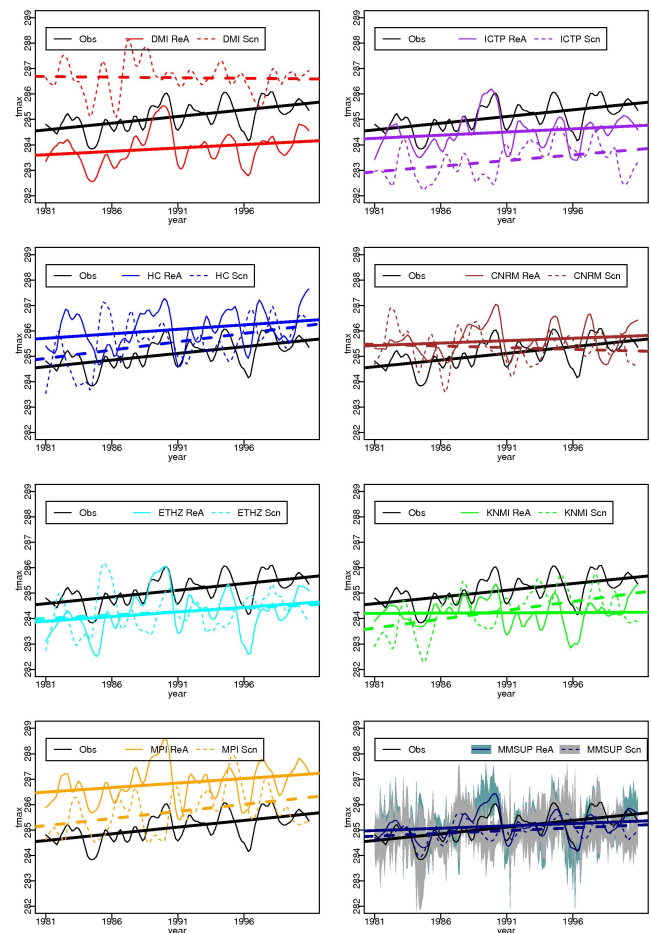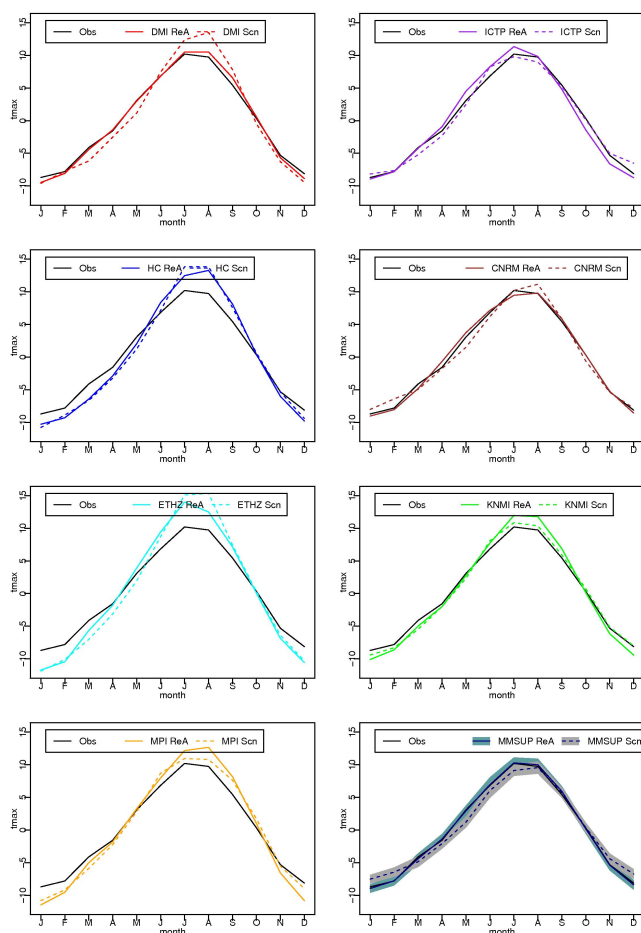

Fig. 5. Comparison between trends from observations obtained with optimal interpolation of Piemonte data (black lines), reanalysis runs (solid lines) and scenario runs (dashed lines) for different models (acronyms in Table 1) and multimodel superensemble (MMSUP) in the period 1981–2000. Multimodel training period: 1961–1980. Multimodel uncertainty (square root of the total multi-normal variance) is represented as confidence bands.

3. the introduction of a "polling" can interrupt dry and wet day continuous series, then the consecutive wet days number and consecutive dry days number have to be checked;

4. this technique does not ensures automatically the comprehensive bias correction, because in this case the correction is "forecast depending": the multimodel final bias has to be checked on annual and short-time ranges.

3 Results

3.1 High resolution data in Piedmont

We tested the technique on the past data, splitting the control period of the models into two halves: the first one

Fig. 6. Comparison between seasonal component from observations obtained with optimal interpolation of Piemonte data (black lines), reanalysis runs (solid lines) and scenario runs (dashed lines) for different models (acronyms in Table 1) and multimodel superensemble (MMSUP) in the period 1981–2000. Multimodel training period: 1961–1980. Multimodel uncertainty (square root of the total multinormal variance) is represented as confidence bands.

(1961–1980) was used as training period, the second one (1981–2000) as forecast period.

We decomposed the models and multimodel time series in the trends and seasonal components with the Seasonal Decomposition of Time Series by LOESS (Cleveland et al., 1990) and compared them with the observation series.

The multimodel superensemble temperature fields show a very good reduction of model biases (Fig. 5) and a very close reproduction of the temperature monthly statistics (Fig. 6). In this paper we show only the validation results of the maximum temperature, but those of minimum temperature have identical skill. Please notice that, in the control period, the reanalyses and scenario runs from the models show not only strong biases towards the observed temperature but, more worrying, the trends sometime differ in a very significant way, and the reanalysis and scenario runs from the same model very often show a different behaviour.

The results of the multimodel superensemble seem to be quite consistent in term of average values, although we cannot assure that our ensemble is able to catch all the variability of the underlying climate, because the number of ensemble members is limited and they are "best effort" simulations, trying to asses a realistic behaviour of the climate, rather than a perturbed physics ensemble trying to span the uncertainties of the modelling system (according to the terminology by Knutti et al., 2010).

The Walter and Lieth (1960–1967) diagrams referred to precipitation produced by climate models show very strong biases (up to 200 % during winter months) in the Alpine region when compared with observations. In Fig. 7 we compare the Walter and Lieth diagrams after removing each model yearly averaged bias, to obtain a fair comparison with multimodel which is almost unbiased. Multimodel does not show very large biases in any month and reproduces the precipitation annual distribution quite well, both in time and amount. Only two input models out of seven have quite comparable skill, not taking into account their large average biases.

Nevertheless, the multimodel post-processing of precipitation allows correcting the statistical properties of the models to reduce the strong models biases, to reproduce the correct precipitation monthly statistics and the average number of consecutive dry periods (more than 5 days without precipitation, namely < 1 mm), Fig. 8.

On the other hand, it is less effective in reproducing the observed average number of consecutive wet periods defined as more than 5 days with precipitation larger than 1 mm. Being that in the Southern Alps the probability of having a dry day (and therefore to extract a dry day and interrupt a wet days series) is much higher than the probability of having a wet day (and then to interrupt a dry day series), the probabilistic sampling from the multimodel PDF can introduce a gap in a continuous series of wet days. We are evaluating a technique to avoid this problem by substituting the "white noise" random number generation used to extract the values from the distribution with a function describing the correct correlation between the consecutive days of rain, but this work is still ongoing.

Figures 9–11 show the difference between the multimodel superensemble scenario data averaged over the period 2031–2050 with respect to the period 1981–2000, as a function of the season (comparison was made on the scenario for better consistency, but the scenario is very close to observations). The scenario projection shows a significant increase of the temperatures over the region. This increase is shown also by the original RCMs, but the post-processed data allow a better characterization of the Alpine region, with an increasing and more realistic variance of temperature variations as a function of the altitude, thanks to the calibration with observations.

In particular, maximum temperatures averaged in the study area show a significant increase in winter ($+0.8 \,^\circ$C), spring ($+1.4 \,^\circ$C), summer ($+1.6 \,^\circ$C) and autumn ($+1.2 \,^\circ$C limited
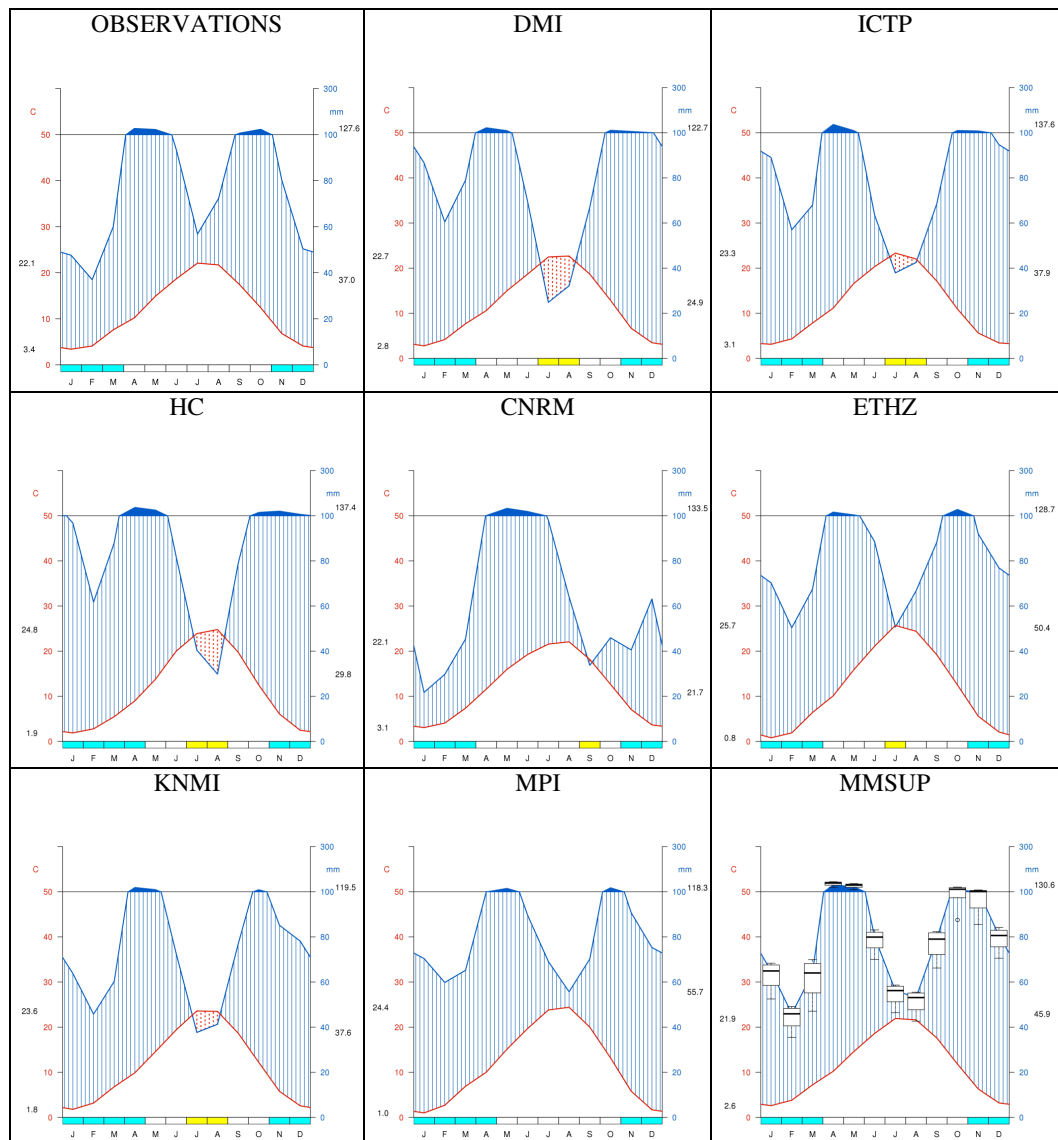
Fig. 7. Walter and Lieth diagrams of the models and multimodel (MMSUP) for the values averaged over Piedmont OI grid points, period 1981–2000. The yearly averaged bias was subtracted from each monthly value. Multimodel uncertainty is represented as boxplots of the precipitation in the MonteCarlo experiment (5th, 25th, 50th, 75th, 95th percentiles are shown).

to the mountains). Maximum temperatures during spring and summer increase more on the plains than in the mountains.

Minimum temperatures show a significant increase in winter ($+1.1\,°C$), spring ($+1.3\,°C$), summer ($+1.8\,°C$) and autumn ($+1.3\,°C$ limited to the mountains). Minimum temperatures during autumn and winter increase more on the plains than in the mountains.

Precipitations on the annual scale show a slight decrease (not statistically significant with 95 % confidence level), while on a seasonal basis they show a significant decrease in spring ($-9\,\mathrm{mm\,month^{-1}}$ only in the western Alps), summer ($-22\,\mathrm{mm\,month^{-1}}$), with few differences among mountains and plains and in autumn ($-26\,\mathrm{mm\,month^{-1}}$ limited to the mountains).

## 3.2 The greater Alpine area perspective

To enlarge our perspective, we applied the same techniques to the whole Alpine areas included in the GAR region. As mentioned above, the reference observations were extracted from the E-OBS dataset, and a preliminary comparison was made with the higher resolution gridded dataset covering Piedmont region to check their compatibility over the common geographical area. Very few meteorological stations used to produce the two dataset are in common, so they can be considered independent each other.

The two datasets agree very well in the trends of maximum and minimum temperatures and also maximum temperature absolute value, while the OI minimum temperatures
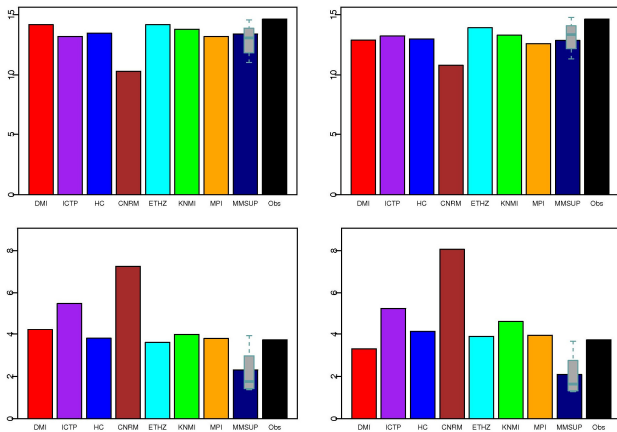
Fig. 8. (top) Number of dry periods (5 consecutive days with precipitation < 1 mm)/year for reanalysis (left) and scenario (right); (bottom) number of wet periods (5 consecutive days with precipitation > 1 mm)/year for reanalysis (left) and scenario (right); input models (colours), multimodel (blue) and observations (black), period 1981–2000. Precipitation is calculated as the average over the Piedmont grid points. Multimodel uncertainty is represented as boxplots of the dry/wet mean number of days in the MonteCarlo experiment (5th, 25th, 50th, 75th, 95th percentiles are shown).
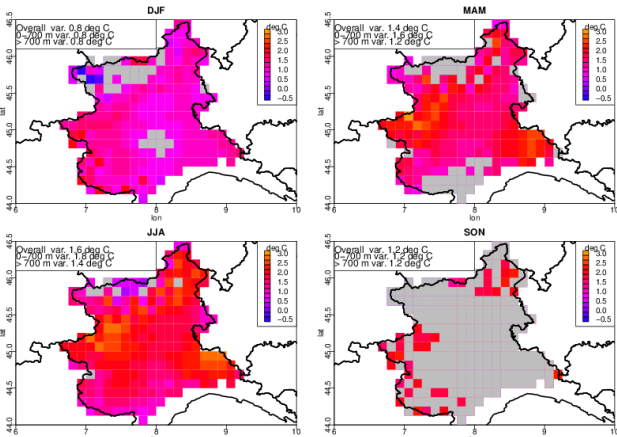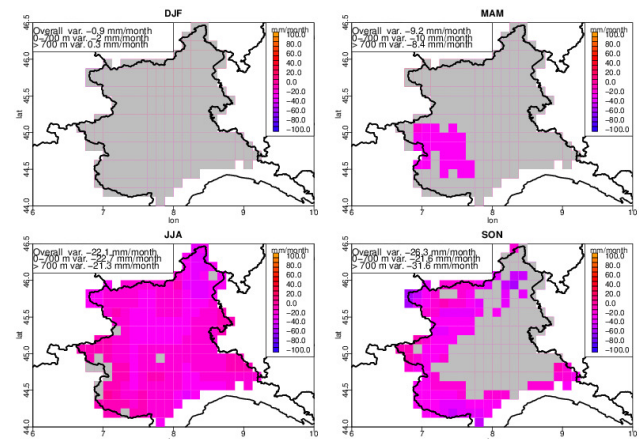

Fig. 9. Difference between the multimodel superensemble scenario maximum temperatures averaged over the period 2031–2050 with respect to the period 1981–2000, as a function of the season (*T* test conf. level 95 %) in Piemonte region. In the upper left boxes overall averages over significant points and altitude bands averages are shown.

are warmer than the E-OBS ones by almost 1 °C. The differences can be explained only partially with a different average elevation of the two datasets, and will be the object of a further investigation.

Anyway, we are mainly interested in variations rather than absolute values, then the strong agreement between the dataset trends in temperature allows for a comparison.
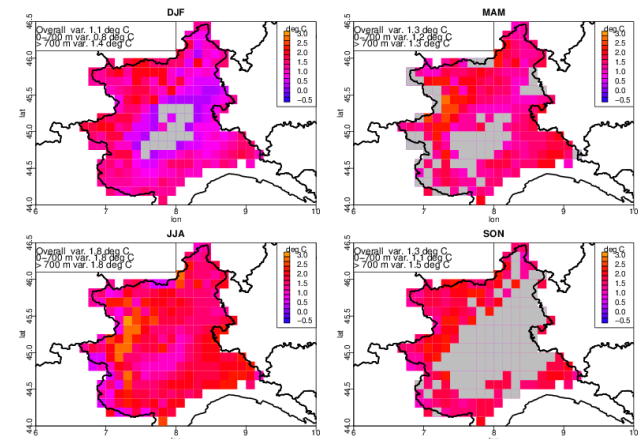
Fig. 10. Difference between the multimodel superensemble scenario minimum temperatures averaged over the period 2031–2050 with respect to the period 1981–2000, as a function of the season (*T* test conf. level 95 %) in Piemonte region. In the upper left boxes overall averages over significant points and altitude bands averages are shown.
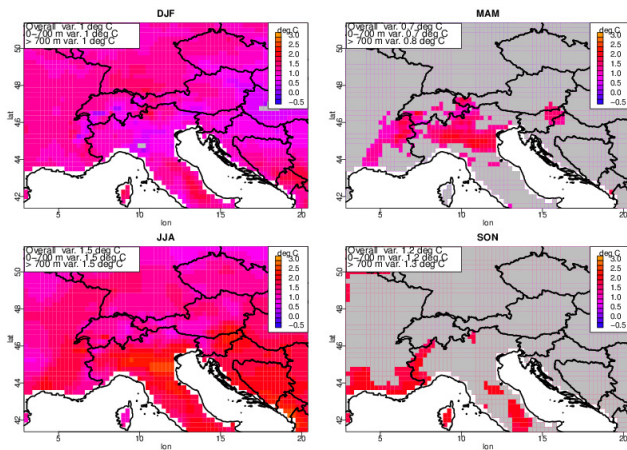

Fig. 11. Difference between the multimodel superensemble scenario precipitation averaged over the period 2031–2050 with respect to the period 1981–2000, as a function of the season (*T* test conf. level 95 %) in Piemonte region. In the upper left boxes overall averages over significant points and altitude bands averages are shown.
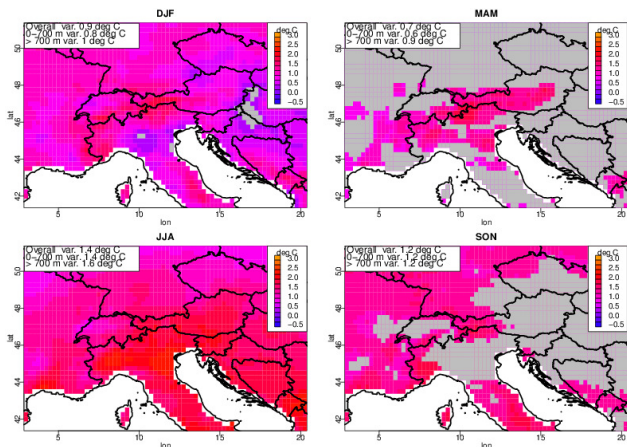
Precipitations agree quite reasonably if we take into account the average over the whole common area, but when we look at specific points in the higher mountains, the climate regime described by the two dataset is quite different, then the comparison is more difficult.

First of all, we repeated the validation of the multimodel results in the control period, with the same agreement already shown for the high resolution dataset.

Figures 12–14 show the difference between the multimodel superensemble scenario data averaged over the period 2031–2050 with respect to the period 1981–2000, as a function of the season.

Fig. 12. Difference between the multimodel superensemble scenario maximum temperatures averaged over the period 2031–2050 with respect to the period 1981–2000, as a function of the season (T test conf. level 95 %) in the GAR area.

Fig. 14. Difference between the multimodel superensemble scenario precipitation averaged over the period 2031–2050 with respect to the period 1981–2000, as a function of the season (T test conf. level 95 %) in the GAR area.
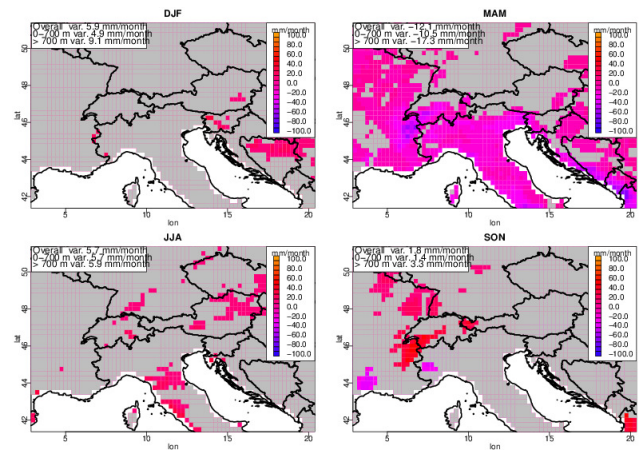
The decrease shown by the GAR data is quite light compared to the one described by the high resolution dataset. We cannot then drive a definitive conclusion about the precipitation behaviour from these two different evaluations.

3.3 Conclusions and future developments

Multimodel techniques can be used fruitfully to better evaluate the climatic parameters in complex orography regions. Multimodel superensemble provides a good estimation of temperature and data in Piedmont, with a very good reduction of the biases and a good reproduction of the monthly variations. We introduced here the first application of a new probabilistic multimodel superensemble dressing to precipitation, providing a reasonably good estimation of the precipitation regime in Piedmont.

We evaluated and validated the multimodel results on two independent datasets, the E-OBS dataset and an high resolution optimal interpolation of the Piedmont station data.

Regarding the common geographical area of the two calibration datasets (Piedmont), the temperatures show similar behaviour in the mid-XXI century scenario, with a general increase compared with the control period, significant in all the seasons except for autumn. The OI data show stronger increases in the higher elevation, while the E-OBS data have the same signal with no elevation dependence.

On the other hand, precipitation variations in the scenario depend more on the observations used for the calculation of the multimodel. In both calibrations, precipitation is not projected to change significantly at an annual scale, while at a seasonal scale we found a decrease in summer precipitation as regards the OI dataset and in spring about the E-OBS dataset.

Fig. 13. Difference between the multimodel superensemble scenario minimum temperatures averaged over the period 2031–2050 with respect to the period 1981–2000, as a function of the season (T test conf. level 95 %) in the GAR area.

The minimum and maximum temperatures show an increasing trend everywhere, not always significant, in particular in spring and autumn. During summer and winter the major increases occur in the southern part of the dominium, with temperature increases up to 2 °C/50 yr in the Po valley.

If we compare the results in Piedmont with the result from the OI, we can see that the GAR signal is similar but weaker and flatter, with less emphasis on the Alpine chain.

As in the OI case, the annual precipitation average over the GAR area does not vary in a significant way, and a significant decrease can be seen only in spring. The timing of the precipitation decrease differs from that observed on the OI grid, and this mismatch can arise from the different precipitation regimes as seen by the two observation datasets.

As concerns the Greater Alpine Region, the projection to the mid-XXI century shows a quite uniform temperature increase between plain and mountain regions in all the seasons, except than in spring, when the increase is significant only in the mountains. Precipitation does not show any significant annual variation, and on a seasonal basis it shows a significant decrease in spring only.

We are evaluating a technique to better describe the correlation of the daily precipitation and to allow a more correct random extraction of a given day from the multimodel-corrected PDF.

We will also test the application of the new multimodel superensemble dressing technique to temperature data, in order to allow a comparison with the standard techniques.

Several impact studies are ongoing with the use of these data, about mountain hydrology, wildfire potential, permafrost, Alpine lakes biology, mountain biodiversity, heat waves.

References

Cane, D. and Milelli, M.: Weather forecasts obtained with a Multimodel SuperEnsemble Technique in a complex orography region, Meteorol. Z., 15, 207–214, 2006.

Cane, D. and Milelli, M.: Can a Multimodel SuperEnsemble technique be used for precipitation forecasts?, Adv. Geosci., 25, 17–22, 2010,
http://www.adv-geosci.net/25/17/2010/.

Cane, D., Wastl, C., Barbarino, S., Renier, L., Schunk, C., and Menzel, A.: Projection of fire potential to future climate scenarios in the Alpine area: some methodological considerations, Climatic Change, online first, doi:10.1007/s10584-013-0775-7, 2013a.

Cane, D., Ghigo, S., Rabuffetti, D., and Milelli, M.: Real-time flood forecasting coupling different postprocessing techniques of precipitation forecast ensembles with a distributed hydrological model. The case study of may 2008 flood in western Piemonte, Italy, Nat. Hazards Earth Syst. Sci., 13, 211–220, doi:10.5194/nhess-13-211-2013, 2013b.

Christensen, J. H., Kjellström, E., Giorgi, F., Lenderink, G., and Rummukainen, M.: Weight assignment in regional climate models, Clim. Res., 44, 179–194, 2010.

Ciccarelli, N., von Hardenberg, J., Provenzale, A., Ronchi, C., Vargiu, A., and Pelosini, R.: Climate variability in north-western Italy during the second half of the 20th century, Global Planet. Change, 63, 185–195, 2008.

Cleveland, R. B., Cleveland, W. S., McRae, J. E., and Terpenning, I.: STL: A seasonal-trend decomposition procedure based on loess, Journal of Official Statistics, 6, 3–73, 1990.

Coppola, E., Giorgi, F., Rauscher, S.-A., and Piani, C.: Model weighting based on mesoscale structures in precipitation and temperature in an ensemble of regional climate models, Clim. Res., 44, 121–134, 2010.

Feng, J., Lee, D.-K., Fu, C., Tang, J., Sato, Y., Kato, H., Mcgregor, J. L., and Mabuchi, K.: Comparison of four ensemble methods combining regional climate simulations over Asia, Meteorol. Atmos. Phys., 111, 41–53, 2011.

Giorgi, F. and Mearns L. O.: Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the "reliability ensemble averaging" (REA) method, J. Climate, 15, 1141–1158, 2002.

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M.: A European daily high-resolution gridded dataset of surface temperature and precipitation, J. Geophys. Res. Atmos., 113, D20119, doi:10.1029/2008JD010201, 2008.

Kalnay, E.: Atmospheric modeling, data assimilation and predictability, Cambridge Univ. Press, 341 pp., 2003.

Knutti, R., Abramowitz, G., Collins, M., Eyring, V., Gleckler, P. J., Hewitson, B., and Mearns, L.: Good Practice Guidance Paper on Assessing and Combining Multi Model Climate Projections, in: Meeting Report of the Intergovernmental Panel on Climate Change Expert Meeting on Assessing and Combining Multi Model Climate Projections, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., and Midgley, P. M., IPCC Working Group I Technical Support Unit, University of Bern, Bern, Switzerland, 2010.

Krishnamurti, T. N., Kishtawal, C. M., LaRow, T. E., Bachiochi, D. R., Zhang, Z., Williford, C. E., Gadgil, S., and Surendran, S.: Improved weather and seasonal climate forecasts from Multimodel SuperEnsemble, Science, 285, 1548–1550, 1999.

Krishnamurti, T. N., Kishtawal, C. M., LaRow, T., Bachiochi, D., Zhang, Z., Williford, C. E., Gadgil, S., and Surendran, S.: Multimodel superensemble forecasts for weather and seasonal climate, J. Climate, 13, 4196—4216, 2000.

Nicolella, M. and Cane, D.: Indices bio-météorologiques en Piémont: évaluation de l'impact du changement climatique avec down-scaling de modèles climatiques régionaux, in: Les climats régionaux: observation et modélisation, edited by: Bigot, S. and Rome, S., ISBN 978-2-907696-18-0, 2012.

Palmer, T. N., Alessandri, A., Andersen, U., Cantelaube, P., Davey, M., Délécluse, P., Dequé, M., Diéz, E., Doblas-Reyes, F. J., Feddersen, H., Graham, R., Gualdi, S., Guérémy, J.-F., Hagedorn, R., Hoshen, M., Keenlyside, N., Latif, M., Lazar, A., Maisonnave, E., Marletto, V., Morse, A. P., Orfila, B., Rogel, P., Terres, J.-M., and Thomson, M. C: Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (DEMETER), B. Am. Meteorol. Soc., 85, 853–872, 2004.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T.: Numerical Recipes in Fortran, Cambridge University Press, Cambridge, 256 pp., 1992.

Raftery, A. E., Balabdaoui, F., Gneiting, T., and Polakowski, M.: Using Bayesian averaging to calibrate forecast ensembles, Tech. Rep. 440, Dep. of Stat., Univ. of Wash., Seattle, 2003.

Raftery, A. E., Gneiting, T., Balabdaoui, F., and Polakowski, M.: Using Bayesian model averaging to calibrate forecast ensembles, Mon. Weather Rev., 133, 1155–1174, 2005.

Randall, D. A., Wood, R. A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., Stouffer, R. J., Sumi, A., and Taylor, K. E.: Climate Models and Their Evaluation, in: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

Todini, E.: A model conditional processor to assess predictive uncertainty in flood forecasting, Int. J. River Basin Manage., 6, 123–137, 2008.

Yun, W. T., Stefanova, L., and Krishnamurti, T. N.: Improvement of the multimodel superensemble technique for seasonal forecasts, J. Climate, 16, 3834–3840, 2003.

Walter, H. and Lieth, H.: Klimadiagramm-wetatlas, G.Fisher, Jena, 1960–1967.

Weigel, A. P., Knutti, R., Liniger, M. A., and Appenzeller, C.: Risks of model weighting in multimodel climate projections, J. Climate, 23, 4175–4191, 2010.