Probabilistic correction of RCM precipitation in the Basque Country (Northern Spain)

Robert Monjo*[,1], Guillem Chust[2] and Vicente Caselles[3]

[1]Fundación para la Investigación del Clima, C/ Tremps 11, Madrid 28040 (Madrid, Spain)
[2]AZTI-Tecnalia, Marine Research Division, Txatxarramendi ugartea z/g, 48395 Sukarrieta (Vizcaya, Spain)
[3]Department of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia. N. 50, Dr. Moliner. 46100 Burjassot (Valencia, Spain)

*Corresponding author: E-mail: robert@temps.cat

ABSTRACT

A parametric quantile-quantile transformation is used to correct the systematic errors of precipitation projected by regional climate models. For this purpose, we used two new probability distributions: modified versions of the Gumbel and Log-Logistic distributions, which fit to the precipitation of both wet and dry days. With these tools, the daily probability distribution of 7 regional climate models (RCM) was corrected: Aladin-ARPEGE, CLM-HadCM3Q0, HIRHAM-HadCM3Q0, HIRHAM-BCM, RECMO-ECHAM5-rt3, REMO-ECHAM-rt3 and PROMES-HadCM3Q0. The implemented method presents an error less than 5% on the simulation of the average precipitation, and 1% on the simulation of the number of dry days. For the study area, an intensification of daily and subdaily precipitation is expected under the A1B scenario, throughout the 21st century. This intensification is interpreted as a consequence of the process of 'mediterraneanisation' of the most southern ocean climate.

Key words: Probabilistic correction, Bias correction, Empirical downscaling, Transfer functions, Precipitation, Model Output Statistics

1. Introduction

Water is the most important natural resource for most human activities such as agriculture and industry. But sometimes, the unevenly distribution of water can also result in a risk of flooding events that may affect the urban environment and socio-economic activities (Olcina 2008). Accordingly, the study of precipitation in the context of climate change is critical for a land-use planning and economic activities.

In order to analyse the potential climate changes in precipitation regime, numerical models try to simulate the climatic conditions of the global atmospheric circulation. Regional Climate Models (RCM) shows advantages for some regions of world, especially with little sharp relief. However, for hilly areas, a low spatial resolution causes a weak correspondence between the climate simulated for each grid point and the observed one in the stations located at subgrid scales. Hence, greater regionalization procedures is needed for capture the climatic change signals of the RCMs in relation to the climatic features at local level (Widmann and Bretherton 2003; Piani et al. 2010).

A way for regionalize the climate change signal is the calibration of models using empirical data from stations. In particular, it is possible by using statistical correspondence between the observed and simulated ECDF (Empirical Cumulative Distribution Functions). This correspondence is applied with several variants: quantile mapping (Déqué 2007; Formayer and Haas 2010), bias correction (Ines and Hansen 2006; Piani et al. 2010), transfer functions (Benestad 2010; Kallache et al. 2011), empirical-statistical downscaling (Maraun et al. 2010), probabilistic downscaling (Michelangeli et al. 2009), Model Output Statistics (Maraun et al. 2010; Turco et al 2011), etc.

Comparing with other methods of regionalization, the probability correspondence or qualtile-mapping shows good results to correct the direct output of the RCM (Themeβl et al. 2010). However, this method can show overfitting because the ECDF have often irregularities in the continuity of the curve. These irregularities may be due to natural variability of low-occurrence events or systematic errors of the measuring instrument. The natural contribution to the irregularities can be smoothed for very long time-series.

Another problem is that empirical quantile-quantile transformation of highest values of precipitation has a large range of error. In addition, they cannot be extrapolated to longer periods of time beyond the actual length of the reference series (Piani et al. 2010). In this sense, it seems that the empirical quantile-mapping cannot be used to correct the heaviest rainfall in climate models because in general the series of stations are much shorter.

All these problems can be solved in part if the quantile-mapping is applied using mathematical functions that fit robustly to the probability distribution of precipitation. Thus, extrapolation of extreme precipitation has a minor error because the irregularities of ECDF are eliminated. In this line of work, some authors have used the gamma distribution to fit to the precipitation from wet days (Watterson and Dix 2003; Watterson 2005; Elshamy et al. 2009; Piani et al. 2010), but in this paper we propose to use other alternatives, valid for some climate regions.

Therefore, the main objective of this work is to apply two new probability distributions to correct the bias of different projections of RCM precipitation, of both wet and dry days. In particular, we have applied the bias correction for the Basque Country (Northern Spain), for this century under the A1B scenario. Furthermore, this work seeks to achieve an appropriate time-structure of simulated rainfall, which is required for hydrology applications in this region (Mendizabal et al. 2013).

The selected study area is also climatically interesting because the north of the Iberian Peninsula lies in a borderland between the decrease (south) and the increase in rainfall (north), according to most climate projections for the twenty-first century (Christensen et al. 2007, Goubanova and Li 2007). The expansion of the Hadley cell could push the tropical subsidence zone (desert areas) up into the southern peninsula (Lu et al. 2007). At the same time, it is expected that air warming will facilitate an increase in precipitable water content in Northern Europe, causing a greater intensity of rainfall (Déqué et al. 2007, Goubanova and Li 2007).

2. Materials

2.1. Study area

The Basque Country (Northern Spain) is geographically characterised by an important mountainous relief (Figure 1a). The coastal relief can reach locally up to 1,000 m, and it is responsible for the formation of a rugged coastline, with vertical cliffs intercalated by small estuaries. Thus, the rivers draining to the Basque coast

are torrential, with very short time-lags between the precipitation and resulting river discharge (Uriarte et al. 2004).

The Basque Country is characterised primarily by an Atlantic climate, over in the remainder of the Northern Iberian Peninsula. Due to its location, this area is predominantly affected by the polar jet waves or Rossby waves, but sometimes also by the subtropical jet (Alves and Verdière 1999; Peliz et al. 2002). The orographic features of this region cause a sharp contrast between the northern and southern rainfall (Garmendia et al. 1989). In fact, annual precipitation ranges from 400 mm in the south of Alava, with Mediterranean climate (Martín-Vide 2004), and 2500 mm in the northeast of Guipúzcoa, with oceanic climate (Capel-Molina 2000).

2.2. Data

The data of 67 stations of State Meteorological Agency of Spain (AEMET) have been used with daily availability (Fig. 1a). However, many stations present data gaps for a few years. Stations show an average of 15% of gaps respect to each series length. For the reference period 1961-2000 only 11 stations have full series, while the total data of 67 stations is just over 70% compared to the maximum possible dataset. Therefore, it was necessary to use a cross reference series of precipitation in the Basque Country, spatially reconstructed using a digital terrain model and a multi-linear fit (Moncho et al. 2009b).

The models used in this study are the numerical outputs of 7 RCMs, available from the European project ENSEMBLES-rt3 (http://ensemblesrt3.dmi.dk/). These models have a daily temporal resolution and spatial resolution of 25 km × 25 km, and are forced with the SRES-A1B future scenario, corresponding to an intermediate stage between different projections of increased greenhouse gases (Christensen et al. 2007, Meehl et al. 2007, Niehörster et al. 2008). The used RCMs were:

a) Aladin-ARPEGE: National Center for Meteorological Studies (CNRM), France.
b) CLM-HadCM3Q0: Swiss Institute of Technology, Department of Geography (ETHZ), Switzerland.
c) HIRHAM-HadCM3Q0: Norwegian Meteorological Institute (METNO).
d) HIRHAM-BCM: Norwegian Meteorological Institute (METNO).
f) RECMO-ECHAM5-rt3: Meteorological Institute Netherlands (KNMI).
g) REMO-ECHAM-rt3; Max-Planck Institute (MPI), Germany.
h) PROMES-HadCM3Q0: University of Castilla–La Mancha (UCLM), Spain.

All RCM outputs provided time-series until 2100 except PROMES-HadCM3Q0, whose projection ends in 2040. The topography of a small region like the Basque Country is highly smoothed in the RCMs (Fig. 1b): For example, the highest peak considered in the grid is around 800 m, compared with almost 1600 m of Aitxuri Mt. (Guipúzcoa). However, it is expected that climate change signals be reflected somehow in the change of precipitation probability distribution, both for each grid point and the spatial average of the studied area. In particular, we use the average of the grid points corresponding primarily to the Basque Country (18 ground points).

3. Methodology

3.1. Probability distribution functions

The cumulative probability $\pi(p \geq P)$ is defined as the probability that a station records an event with rainfall p equal to or greater than P in one day, i.e, it is the

cumulative frequency with which it happens. The return period or the expected time of this event is given by the inverse of the cumulative probability, $1/\pi(p \geq P)$.

In order to model the Cumulative Distribution Function (CDF) of daily precipitation, we used a modified version of Gumbel distribution and another of Log-Logistic (Moncho et al. 2012), respectly:

$$\pi(\lambda_0 \geq \lambda) = \exp\left[-\exp\left(\lambda^w + k\right)\right] \quad (1)$$

$$\pi(\lambda_0 \geq \lambda) = \frac{1}{1+\lambda^{w+\lambda^w e^{-k}}} \quad (2)$$

where w and k are fitting parameters; λ is the relative precipitation, which also depends on two fitting parameters: location parameter, P_o, which represents the most probable precipitation, and the other one is the scale parameter, P_1:

$$\lambda \equiv \frac{P - P_o}{P_1} \quad (3)$$

The most likely value for each parameter was estimated by statistical inference, in particular with the Profile Log-Likelihood (Akaike 1975, Raue et al. 2009). A standardisation was performed by correspondence between the values of precipitation and its cumulative probability (CDF). Consequently, time-series of daily precipitation can be transformed in a time-series of return period, where each of these return periods is associated with one value of daily precipitation. The resulting time-series will be called Series of Return Periods or SRP. Partial standardisation is defined as the process of generating a SRP by fitting a CDF to a portion of the total length of the series (for example, a reference time-period).

In order to measure the goodness-of-fit for the general and extreme values, both fitted functions (Eq.1 and 2) are tested using two nonparametric tests: Kolmogorov-Smirnov test (KS, Marsaglia et al. 2003, Sekhon 2010) and Anderson-Darling test (AD, Scholz and 1987), which is most sensitive to the tails. Fits with p-values lower than 0.05 are rejected.

3.2. Gap-filling and probabilistic correction

A great length of observed time-series is required to compare them climatically with the simulated ones by the models. Therefore, gap-filling and extension of observed time-series are necessary for a good performance of the bias correction. The same approach of probabilistic transformation is used for both the gap-filling and the bias correction. It is a parametric quantile-mapping method, i.e., it consists in transferring the shape of one CDF to another (Benestad 2010, Kallache et al. 2011).

A generic two-step algorithm is used for both probabilistic transformations. The first step is to obtain the SRP of each time-series using partial standardisation in the common period between the ‘emitter series’ (which transfers the CDF shape) and the ‘receptive series’ (which receives it). For the gap-filling, the ‘emitter series’ are the shortest observed ones (less than 40 years) and the ‘receptive series’ are the longest observed series (at least 40 years). For the probability correction, the ‘emitter series’ are the extended observed time-series, and the ‘receptive series’ are the time-series simulated by the models. Each SRP is obtained using the theoretical probability distribution (Eq. 1 or 2) that obtained the lowest error for each fit. In particular, the goodness-of-fit is given by the mean normalized absolute error (MNAE), calculated over all values of each CDF.

The second step consists in transferring climate features of each ‘emitter series’ to the SRP of each ‘receptive series”, applying the parameters fitted to the CDF of the ‘emitter series’ (inverse standardisation). In the gap-filling process, for each ‘emitter SRP’, the most similar ‘receptive SRP’ is selected according to two

statistical criteria: the Pearson correlation (R) and the MNAE. Specifically, we chose the lowest value of (1-R)·MNAE. If the best 'receptive SRP' has a MNAE higher than 20%, the 'emitter series' cannot be successfully extended. Then, it is rejected for the climate analysis of time variability (daily scale) but not for the analysis of climate averages.

3.3. Analysis of mean and extreme precipitation

The corrected time-series of the climate models are analysed to validate the simulated number of precipitation days and the mean and standard deviation of precipitation. The statistical criteria used for the validation are the relative mean absolute error (RMAE), the relative bias (RBIAS) and the Pearson correlation (R) of the spatial distribution. The relative change in precipitation is obtained with a weighted average of the seven corrected RCM projections. Weights are defined as:

$$w_i \equiv \frac{\left(RMAE_i^{\,2} + RBIAS_i^{\,2} + (1-R_i)^2\right)^{-1}}{\sum_{i=1}^{7}\left(RMAE_i^{\,2} + RBIAS_i^{\,2} + (1-R_i)^2\right)^{-1}} \quad (4)$$

where w_i is the weight for the 'i' RCM projection. The relative absolute error of the ensemble average is estimated by the same equation. Using this value, confidence intervals are obtained for the projected change in average precipitation.

Expected precipitation for a return period of 100 years is obtained with the best fitted curve (Eq.1 or Eq. 2), i.e, the curve that obtained the lowest MNAE in the fit to ECDF of each station. Change projected in the 100y-return precipitation for the beginning of this century (2001-2040) comparing with the past (1961-2000) is calculated using the same method of ensemble average that we have described above (Eq. 4).

In order to analyse the performance of the simulated time-structure of the precipitation, properties of wet spells are studied by using the method of index *n* (Moncho et al. 2009a). Given a probabilistic model, it is possible to estimate the expected rainfall using a function of return period, $f(T)$. Hence, the expected precipitation P and P_o, for duration of t and t_o respectively, can be described by a generic curve of Intensity-Duration-Frequency (IDF):

$$\frac{P}{P_o} \approx f(T, T_o)\cdot\left(\frac{t}{t_o}\right)^{1-n} \quad (5)$$

where $f(T, T_o)$ is a function of two return periods (T and T_o) which depends on the chosen probabilistic distribution (inverse of Eq. 1 or 2), and n is the index of precipitation. This index was estimated for all stations and simulations from the precipitation of at least 3 days. This tool was also used to evaluate the simulation time of heavy precipitation events (exceeded 50 mm in one day, in at least three consecutive rainy days). Finally, we studied the possible changes in the IDF along the 21st century, in terms of the *n* index and the length of wet spells.

All data and result treatments were carried out by using statistical packages based on R language (R Development Core Team, 2010), mainly "stats" and "fields" (Furrer 2012).

4. Results

4.1. Extension and gap-filling

The process of extension and gap-filling is successfully completed in 34 time-series shorter than 40 years. Time-series length was extended between 3 and 64% (mean

of 21%) to complete the 40-year reference period (1961-2000). It was found that the Pearson correlation between each observed series and the most similar one ranged from a minimum of 0.61 to a maximum of 0.90, and the mean is 0.74. The mean normalised absolute error (MNAE) of the simulated ECDF was between 0.02 and 0.12 (with mean 0.05) compared with the observed ECDF. In 72% of the cases, the best fitted curve was the generalised version of the log-logistic distribution (Fig. 2). Therefore, the seven RCMs were transformed to simulate the daily precipitation of the 45 rain gauges with at least 40 years of data (11 full series and 34 extended series).

4.2. Validation of probabilistic transformation

If the monthly average and the daily deviation of each simulated series are compared with the corresponding observed series, it can be noted that the relative value of mean absolute error (RMAE) of both is generally between 1 and 15% depending on the season and RCM (Fig. 3). The average value of RMAE for all stations simulated falls below 5% in all cases and in all months, except for the Aladin-ARPEGE for the summer months (which has an average error around 10%). Similar result is obtained for the relative bias (RBIAS) which is also less than 5% in most cases, except for Aladin-ARPEGE model. The simulated number of dry days even has a lower relative error, around 1%, and a negligible mean bias (Fig. 4).

In addition to low relative error of global results, the simulated spatial variability of precipitation average has a high correlation with observed precipitation average. For example, the correlation of monthly average is around R^2 = 0.99 for all models, with a slope ranging between 1.05 and 1.14 depending on the model (Table 1). The correlation is slightly lower for the simulation of the daily deviation and number of days without precipitation, with R^2 values ranging between 0.96 and 0.98, and the slope is closer to 1, between 0.95 and 1.02.

4.3. Analysis of changes in mean precipitation and extreme

According to the seven RCMs studied, the mean rainfall may have a slight decrease in the Basque Country, especially in spring. Particulary, it provides a possible decrease of up to 15% in the south of the Basque Country, for the quarter April-June of 2001-2040 with respect to period 1961-2000. In autumn, the average rainfall may decrease up to 10% in the north. However, the p-value of these projections is considerably less than 0.95. For the period 2041-2080, the projected decrease in precipitation in the south in spring ranges between 10 and 30% with 95% confidence. No significant changes are projected by the ensemble analysis for the mean precipitation of the rest of seasons.

Regarding to the daily precipitation, return levels were analysed for the observed and simulated time-series. In most of cases, best fits were obtained by the modified log-logistic distribution (Fig. 5). In order to analyse possible changes in extreme precipitation, the study focused on the return period of 100 years. For the period 2041-2080, the results show an increase in the intense precipitation of around 30% in most stations and models (Fig. 6). For the period 2001-2040, the variation is not clear, although some western stations in the Basque Country show an increase in extreme precipitation for most of the models studied. In particular, there is an area in the Western Basque Country with a projected change of up to 30%, with an interval of more than 99% (Fig. 7). However, the probability of change is expected to be smaller, since the latter relative error is 15% in the interpolation.

Concerning the index *n*, it was studied the empirical probability distribution for all precipitation and for precipitation that exceeded 50 mm in one day (Fig. 8).

Regarding all precipitation, no significant changes were identified. This is because the type of precipitation is so diverse that it masks any change in the index *n*, at least for the periods studied in the Basque Country. However, for precipitation higher than 50 mm in one day, changes in probability of index *n* are observed (interval between 0.6 and 0.8). In particular, these changes imply an increase of $+0.03 \pm 0.01$ in the average of *n* (from 0.57 to 0.60). Thus, increasing of the maximum daily precipitation (Fig. 7 and 8) is consistent with a greater concentration in the time of rainfalls, for both daily and subdaily scales (due to higher index *n*).

5. Discussion

5.1. Choice of methodology

According to Themeβl et al. (2010), quantile-mapping shows the best performance comparing with other six methods of empirical-statistical downscaling and error correction. Quantile-quantile transformation can be empirical or parametric using theoretical transfer functions (Déqué 2007; Piani et al. 2010; Maraun et al. 2010; Turco et al 2011). However, the commonly used probability distributions do not fit appropriately to very low and very high precipitation at the same time (Begueria 2005, Moncho et al. 2012). For this reason, some authors apply transfer functions only for extreme precipitation (Fowler et al 2010, Kallache et al. 2011), and others apply only for wet days (Ines and Hansen 2006, Piani et al. 2010). We used theoretical curves fitted to the entire precipitation CDF, not only from wet days, and good performance is obtained (Fig 3, 4 and 5). Methods based on transfer functions present an advantage over empirical quantile-mapping. They can incorporate additional irregularities in the probability curve of the corrected CDF of models (overfitting). In addition, theoretical curves can easily be extended to estimate the extreme rainfall, whereas empirical curves cannot. However, there are two possible problems in the use of theoretical fits: seasonality and stationarity.

Seasonality concerns about the difference in the probability of precipitation depending on the season, month or even fortnight. This means that a probability distribution for each season (or lower scales) needs to be fitted so that the time series of return periods (SRP) is completely non-seasonal (no autocorrelation). One of the problems is that we have to decide where we do the cuts to separate the seasonal effect in the probability distributions. In this work, we have chosen the calendar month because twelve sections are considered enough for the resulting SRP to be non-seasonal (Elshamy et al. 2009). In addition, the borders of the cuts are smooth enough, i.e. there is not much difference between the probability of one month and the next one (maybe except July).

Stationarity is related to the fact that the empirical relationships in past variables may be non-stationary in the future. This issue does not affect significantly the probabilistic transformation method for two reasons:

1) When a time series is partially standardised (future with past), we obtain an SRP which contains all the signal of climate change because, in this way, the change of frequency of specific return periods can be analysed: If the intensity of precipitation increases/decreases in the future, an increase/decrease in the frequency of events with a high return period can be observed in the SRP. Therefore, when the CDF parameters (fitted to each observatory) are applied to the SRP of an RCM, it results an automatic (possible) change in the future probability distributions with respect to the reference period (past).

2) Although, for example, September might seem more like August in the future, this possible change is also contained in the SRP. Continuing with this example, it would expect a higher frequency of the lowest return periods (corresponding to

dry days) in September of the future. Then, it would clearly change in the SRP from being non-seasonal in the past to present some seasonal differences in the future.

In short, it appears that the physical links between the probability of precipitation and the climate change scenarios are maintained after standardisation in the form of SRP. This should not be confused with the preservation of the relative change (in percentage) of rainfall before and after probabilistic correction. In fact, the relative change usually not retained by the non-linear transfer functions (Benestad 2010).

Another important aspect of the standardisation process is to use precipitation series with sufficient detail and accurate values for achieving a more flexible processing. For example, the numerical output of a climate model precipitation has many values below 0.1 mm, thus, it is more detailed in the probability distribution than a meteorological station (which in general has a resolution of 0.1 mm).

Finally, note that bias correction of climate models cannot be evaluated directly using cross-validation, because the decadal climate variability simulated by RCMs does not correspond to the observed decadal variability. To evaluate a statistical downscaling method, it would be advisable to use a reanalysis (e.g. ERA40), where a division in shorter time parts is possible for cross-validate (because annual and decadal correspondence exists). Evaluations of the probabilistic correction are performed by other authors (Ines and Hansen 2006; Déqué 2007; Formayer and Haas 2010; Piani et al. 2010; Benestad 2010; Themeβl et al. 2010; Kallache et al. 2011).

5.2. Probabilistic transformation

The probability distributions used in this study were selected because they show better results than others for several climates in Spain (Moncho et al. 2012). For the Basque Country, both distributions have a high goodness-of-fit for the daily rainfall, with a mean error generally less than 6% (Fig. 2). However, the fitting error in summer is higher, due to the high number of days with low precipitation, which makes the relative error shoot up. Furthermore, the used distributions successfully passed both the KS as AD test ($p>0.05$) for 50 of 67 stations at one time. So, fitted distributions are statistically indistinguishable from the observed distribution of these 50 stations. A subset of 45 of these stations could be successfully extended to have at least 40 years of data. The remaining time-series were too short to being able to properly pass the whole extending process and they were rejected for the probabilistic correction of the RCMs.

Before the probabilistic correction, direct output of the RCM softens considerably the moderate and intense rainfall, in exchange to overestimate the number of days with rainfall, in more than double. After the correction, simulated time-series adequately resemble observed time-series, e.g. according to the BIAS of rainy days and the RMAE of average rainfall (Fig. 3 and 4). Nevertheless, it does not guarantee that probability distributions are similar enough to say that they are statistically indistinguishable. Therefore, KS and AD tests were applied again, but this time to evaluate simulated time-series, after probabilistic transformation. Results showed that the p-value of most stations was above 0.05 for both tests, and it follows that the simulated CDFs are generally indistinguishable from observed ones. However, nearly half of the stations obtained a p-value less than 0.05 in the summer months. This can be explained by the large number of days without precipitation. That is, the KS and AD test gives much consideration to values with a high frequency of occurrence; hence, small differences in the number of dry days cause that p-values be close to zero.

5.3. Analysis of precipitation change

One of the most sensitive aspects of the SRP methodology is that it keeps the precipitation time-structure almost intact. That is, a probabilistic transformation can change the scale of the precipitation and the number of dry days, but not the disposition of dry days alternating with wet days. Therefore, it is important the fact that the models can reproduce the observed conditions of wet and dry spells.

In this regard, Figure 8 shows that regional models (RCMs) adequately reproduce the climatic characteristics of wet spells for the reference period 1961-2000. In general, the projection for 2001-2040 shows no remarkable changes compared with the reference period. However, there is a slight reduction in the length of wet spells in all quarters except for winter. This reduction, in addition to the slight increase in the peaks in autumn projected for some stations, implies a slight increase in the index *n*, as shown in Figure 7.

The reduction of the wet spells length is greater in summer and autumn. In contrast, total precipitation of wet spells decreases more clearly in summer than in autumn, in most of stations. This happens because in summer months, rainfall intensity remains approximately constant (Fig. 9).

As for dry spells, the average length is adequately simulated for the reference period (Fig. 10). However, probability distributions of the length of dry spells have certain differences compared with the observations, although it does not affect the seasonal cycle. For the period 2001-2040, numerical outputs of RCMs projected a clear increase in the length of dry spells in spring and summer of more than 0.5 days per period without rainfall. Note that the increased length of dry spells is consistent with the slight decrease in the length of wet spells projected for 2001-2040 (Fig. 9).

The increase in heavy precipitation in the Basque Country can be related to the possible mediterraneanisation of climate (Moreno 2005). The increase in the number of dry days (Fig. 10) and the highest concentration of precipitation in a short time (Fig. 9) support this hypothesis. The possible mediterraneanisation of Basque climate may be driven by the expansion of the Hadley cell (Lu et al. 2007). The global warming projected under the A1B scenario causes a dilation of the cell and, consequently, the subtropical subsidence area would move towards the North. The main physical implications of this phenomenon are two*:* 1) lower precipitation in spring due to the increased presence of anticyclones. 2) Increased intensity of autumn precipitation in the Mediterranean climate area due to increased summer warming and, hence, more water available in the atmosphere.

6. Conclusions

Both the modified Gumbel distribution and the generalised version of the log-logistic distribution have a high goodness-of-fit for the empirical probability of daily precipitation in the Basque Country. In particular, the latter theoretical distribution showed slight better results, with a mean error of 4% compared to 5% obtained by Gumbel distribution. The probabilistic transformation of RCM adequately corrects the probability distribution and makes it indistinguishable from the stations in the reference period. In fact, the errors of the simulated series are generally less than 5% for both average and standard deviation. The error in the number of days without precipitation is less than 1% in most cases. Moreover, this transformation does not directly affect the climate signal of the models, although brings it in the line with the frequency of precipitation probability of the stations. Therefore, the methodology is useful to quantify possible changes in local climate in terms of absolute amount of precipitation and the way it is distributed throughout time (e.g., dry spells, wet spells, index *n*).

The results of this study show that changes in the time distribution of rainfall are expected during the 21st century and under the A1B scenario. Although duration of consecutive days with rainfall may be shortened, rainfall would be concentrated in shorter time-periods, thus increasing the intensity of the daily maximum in autumn. This idea is reinforced because the index *n* projected for the rest of the century shows a slight increase, from 0.57 to 0.60. With regard to highest precipitation, most models project a clear increase. For example, precipitation with a return period of 100 years will increase to 30% in 2001-2040 compared to 1961-2000, in the Western Basque Country. For the period 2041-2080, the increase may exceed 40% in most of areas. In addition to the intensification of precipitation, mainly in autumn, climate models project an increase in the length of dry spells and a slight decrease in summer precipitation. Therefore, these changes indicate a possible mediterraneanisation of the Basque climate.

Acknowledgments. This work is the last part of the doctoral thesis written in the Department of Earth Physics of the University of Valencia. This work is supported by the Department of Environment, Regional Planning, Agriculture and Fisheries of the Basque Government (K-Egokitzen II project, Etortek Funding Program). Likewise, we acknowledge the State Meteorological Agency of Spain (AEMET) and Hydrographics Confederations of Ebro (CHE) and Júcar (CHJ) for providing the data for this study. In particular, we thank to José Ángel Nuñez, head of the Department of Climatology AEMET delegation in Valencia, and Margarita Martín, AEMET delegate in the Basque Country, for their helpful comments. Finally, it is fair to acknowledge the support of Maddalen Mendizabal (Tecnalia) and especially for raising the issue of probability of daily precipitation.

References

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Automat Cont 19: 716-723

Aleklett K, Hook M, Jakobsson K, Lardelli M, Snowden S, Soderbergh B (2010) The Peak of the Oil Age - Analyzing the world oil production Reference Scenario in World Energy Outlook 2008. Energy Policy 38: 1398-1414

Alves M, Verdière AC (1999) Instability dynamics of a subtropical jet and applications to the Azores Front-Current System-eddy driven mean flow. J Phys Oceanogr 29: 837-864

Begueria S (2005) Uncertainties in partial duration series modelling of extremes related to the choice of the threshold value. J Hydrol 303: 215–230

Benestad RE (2010) Downscaling precipitation extremes. Correction of analog models through PDF predictions. Theor Appl Climatol 100: 1–21

Capel-Molina JJ (2000) El Clima de la Península Ibérica. Ed. Ariel, Barcelona

Carlson WB (2011) The Modeling of World Oil Production Using Sigmoidal Functions. Energy Sources, Part B: Economics, Planning, and Policy 6: 178-186

Christensen JH, Hewitson B, Busuioc A, Chen A, Gao X, Held I, Jones R, Kolli RK, Kwon W-T, Laprise R, Magaña Rueda V, Mearns L, Menéndez CG, Räisäinen J, Rinke A, Sarr A, Whetton YP (2007) Regional Climate Projections". 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 [Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA

Déqué M (2007) Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values, Global and Planetary Change 57: 16–26.

Déqué M, Rowell DP, Lüthi D, Giorgi F, Christensen JH Rockel B, Jacob D, Kjellström E, De Castro M, Van Den Hurk B (2007) An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections. Climatic Change 81: 53-70

Elshamy ME, Seierstad IA, Sorteberg A (2009) Impacts of climate change on Blue Nile flows using bias-corrected GCM scenarios. Hydrol Earth Syst Sci 13: 551–565

Formayer H, Haas P (2010) Correction of RegCM3 model output data using a rank matching approach applied on various meteorological parameters. In Deliverable D3.2 RCM output localization methods (BOKU-coontribution of the FP 6 CECILIA project on http://www.cecilia-eu.org/)Furrer R, Nychka D, Sain S (2010) fields: Tools for spatial data. R package version 6.3. http://CRAN.R-project.org/package=fields

Fowler HJ, Cooley D, Sain SR, Thurston M (2010) Detecting change in UK extreme precipitation using results from the climateprediction.net BBC climate change experiment. Extremes DOI 10.1007/s10687-010-0101-y

Furrer R, Nychka D, Sain, S (2012). fields: Tools for spatial data. R package version 6.6.3. http://CRAN.R-project.org/package=fields

Garmendia MI, Pérez C, Rodríguez C, Garmendia J (1989) Factores determinantes de la precipitación anual en la vertiente cantábrica. Acta Samaltina de Ciencias 67: 113-117, 3 Ref

Goubanova K, Li L (2007) Extremes in temperature and precipitation around the Mediterranean basin in an ensemble of future climate scenario simulations". Global and Planetary Change 57: 27–42

Green PJ, Silverman BW (1994) Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach. Chapman and Hall, London

Ines AVM, Hansen JW (2006) Bias correction of daily GCM rainfall for crop simulation studies. Agr For Met 138: 44–53

Kallache M, Vrac M, Naveau P, Michelangeli P-A (2011) Nonstationary probabilistic downscaling of extreme precipitation, J. Geophys. Res., 116, D05113, doi: 10.1029/2010JD014892

Lu J, Vecchi GA, Reichler T (2007) Expansion of the hadley cell under global warming. Geophys Res Let, 34: L06805

Martín-Vide, J (2004) Spatial distribution of a daily precipitation concentration index in peninsular spain. Int J Clim 24: 959-971

Marsaglia G, Tsang WW, Wang J (2003) Evaluating Kolmogorov's distribution. J Stat Software 8/18

Maraun D, Wetterhall F, Ireson AM, Chandler RE, Kendon EJ, Widmann M, Brienen S, Rust HW, Sauter T, Themeßl M, Venema VKC, Chun KP, Goodess CM, Jones RG, Onof C, Vrac M, Thiele-Eich I (2010) Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. Rev Geophys 48, RG3003, doi:10.1029/2009RG000314

Meehl GA, Stocker TF, Collin WD, Friedlingstein P, Gaye AT, Gregory JM, Kitoh A, Knutti R, Murphy JM, Noda A, Raper SCB, Watterson IG, Weaver AJ, Zhao

Z-C (2007) Global Climate Projections. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate

Mendizabal M, Moncho R, Torp P, Sepúlveda J, Gonzalez-Aparicio I (2013): Uncertainty Analysis on Flood Assessment due to Regional Climate Models. Comprehensive Flood Risk Management. Klijn & Schweckendiek (Eds.). CRC Press, Taylor & Francis Group, London, UK. ISBN 978-0-415-62144-12013

Moncho R, Chust G, Caselles V (2012) Alternative model for precipitation probability distribution: application to Spain. Clim Res 51: 23-33 doi: 10.3354/cr01055

Moncho R, Belda F, Caselles V (2009a) Climatic study of the exponent n of the IDF curves of the Iberian Peninsula. Tethys 6: 3-14

Moncho R, Chust G, Caselles V (2009b) Análisis de la precipitación del País Vasco en el período 1961-2000 mediante reconstrucción espacial. Nimbus 23-24: 149-170Moreno J.M (2005) Principales conclusiones de la evaluación preliminar de los impactos en España por efecto del cambio climático. Oficina Española de Cambio Climático (OECC), Ministerio de Medio Ambiente. Madrid

Niehörster F, Fast I, Huebener H, Cubasch U (2008) The stream one ENSEMBLES projections of future climate change. Ensembles Technical Report nº 3, 2008

Olcina J (2008) Riesgo de inundaciones y ordenación del territorio en España. Documents d'Anàlisi Geogràfica 51: 178-180

Peliz A, Rosa TL, Santos AMP, Pissarra JL (2002) Fronts, jets, and counterflows in the Western Iberian upwelling system. J Mar Syst 35: 61-77

Piani C, Haerter JO, Coppola E (2010) Statistical bias correction for daily precipitation in regionalclimate models over Europe. Theor Appl Climatol, 99:187–192

R Development Core Team (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Raue A, Kreutz C, Maiwald T, Bachmann J, Schilling M, Klingmüller U, Timmer J (2009) Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. Bioinformatics 25: 1923–9

Scholz FW, Stephens MA (1987) K-sample Anderson-Darling Tests, Journal of the American Statistical Association, 399: 918–924

Sekhon JS (2010) Matching: Multivariate and Propensity Score Matching with Balance Optimization. R package version 4.7-11. URL http://CRAN.R-project.org/package=Matching.

Themeβl M-J, Gobiet A, Leuprecht A (2010) Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. Int J Clim, doi: 10.1002/joc.2168.

Uriarte A, Collins M, Cearreta A, Bald J, Evans G (2004) Sediment supply and distribution. Borja, A., Collins, M. (Eds.) Oceanography and Marine Environment of the Basque Country, Elsevier Oceanography Series, 70: 97-131.

Watterson IG, Dix MR (2003) Simulated changes due to global warming in daily precipitation means and extremes and their interpretation using the gamma distribution. J Geophys Res, 108, D13, 4397.

Watterson IG (2005) Simulated changes due to global warming in the variability of precipitation and their interpretation using a gamma-distributed stochastic model. Adv. Water Resourc 28: 1368-1381

Widmann M. and Bretherton C.S (2003) Statistical precipitation downscaling over the Northwestern United States using numerically simulated precipitation as a predictor. J Climate 16: 799-816

TABLES

Table 1. Correlation (R^2) and slope of the climatic values of the average rainfall, daily deviation and dry days in the series simulated by the models, compared with baseline stations.

	Correlation (R^2)			Slope		
RCM	**Mean rainfall**	**Standard deviation**	**Dry days**	**Mean rainfall**	**Standard deviation**	**Dry days**
Aladin-ARPEGE	0.99	0.98	0.97	1.07	0.97	0.99
CLM-HadCM3Q0	0.99	0.96	0.97	1.07	0.95	1.01
HIRHAM-HadCM3Q0	0.99	0.98	0.98	1.14	1.00	0.98
HIRHAM-BCM	0.99	0.98	0.97	1.05	0.97	0.99
RACMO-ECHAM5-rt3	0.99	0.97	0.97	1.08	0.99	1.02
REMO-ECHAM5-rt3	0.99	0.97	0.97	1.08	1.01	0.96
PROMES-HadCM3Q0	0.99	0.97	0.96	1.11	0.97	0.96

Figure captions

Figure 1. a) Location of the stations used in this study (AEMET). b) Example of digital terrain model used in the regional climate models of ENSEMBLES-rt3. c) Election of interest grid points for the Basque Country, which distinguishes between sea points (blue) and ground points (orange).

Figure 2. Comparison of Mean Normalized Absolute Error (MNAE) of the two probabilistic models used for the standardisation of the 45 stations with at least 40 years of data.

Figure 3. Relative mean absolute error (RMAE) and relative bias (RBIAS) of the average and standard deviations of simulated time series. Each bar represents the set of all stations according to each of the seven RCMs.

Figure 4. Absolute and relative mean error (MAE) and bias (BIAS) of the days without precipitation of the simulated series. Each bar represents the set of all stations according to each of the seven RCM.

Figure 5. Return level plots of the daily precipitation for the used stations, in according to the two fitted models: modified Gumbel (red) and modified LogLogistic (blue). Dashed curves correspond to 95% confidence intervals; points correspond to empirical estimates excluding the maximum value of each time-series.

Figure 6. Projection of the expected daily rainfall for a return period of 100 years according to the theoretical curves fitted to the simulated stations. Above shows the projection of the absolute value for the three periods (1961-2000, 2001-2040 and 2041-2080). Below shows the relative change for 2001-2040 and 2041-2080 and the error of the simulation compared with observed time series (1961-2000).

Figure 7. a) Maximum expected rainfall in one day for 1961-2000 simulated with a return period of 100 years. The values are estimated according to the theoretical curves fitted to the simulated stations, and they are interpolated by using a multiple spatial fit. b) Relative change of precipitation described in a), for the period 2001-2040 versus 1961-2000. The black lines represent the confidence intervals estimated from the statistics of the seven RCMs.

Figure 8. a) Empirical probability distribution of the index *n* to the set of all the stations in the period 1961-2000, and the range corresponding to the seven RCM simulations. b) Projection of change for 2001-2040 and 2041-2080. c) and d) Respectively, the same as a) and b) but only the rainfall exceeded 50 mm in one day.

Figure 9. Comparison of observed and simulated wet spells for 1961-2000 and projected for 2001-2040. Each point of the boxplot represents the climatic value of a real or simulated station. Top left shows the average length of spells (days), top right, it shows the average daily peak (in millimeters). Bottom right shows the average total rainfall (in millimeters) and lower left compares the mean index *n* of wet spell.

Figure 10. Evolution of dry spells for the period of 2001-2040 versus the period 1961-2000, both observed and simulated: a) boxplot of the average length of dry spells, and b) cumulative probability of the length of dry spells.

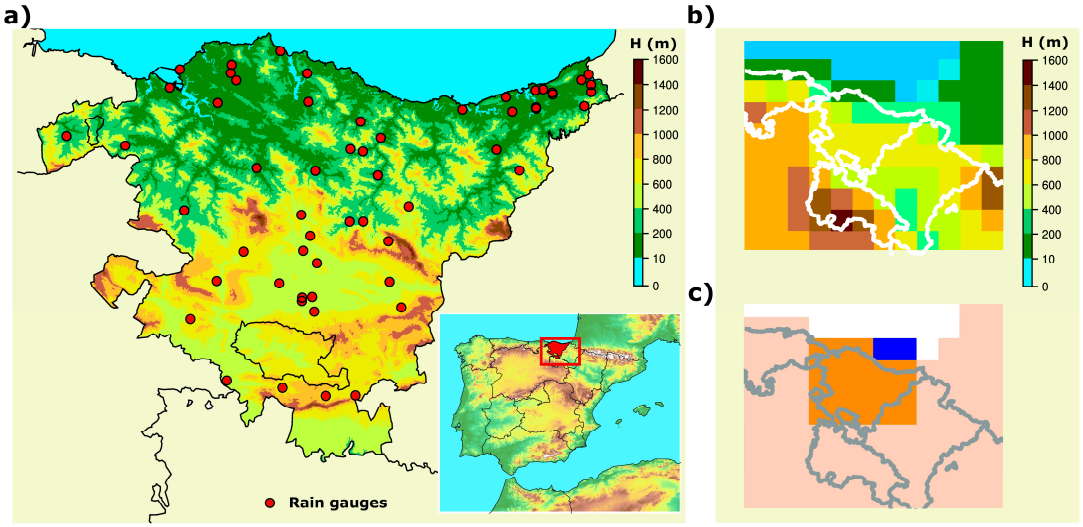

Figure 1. a) Location of the stations used in this study (AEMET). b) Example of digital terrain model used in the regional climate models of ENSEMBLES-rt3. c) Election of interest grid points for the Basque Country, which distinguishes between sea points (blue) and ground points (orange).

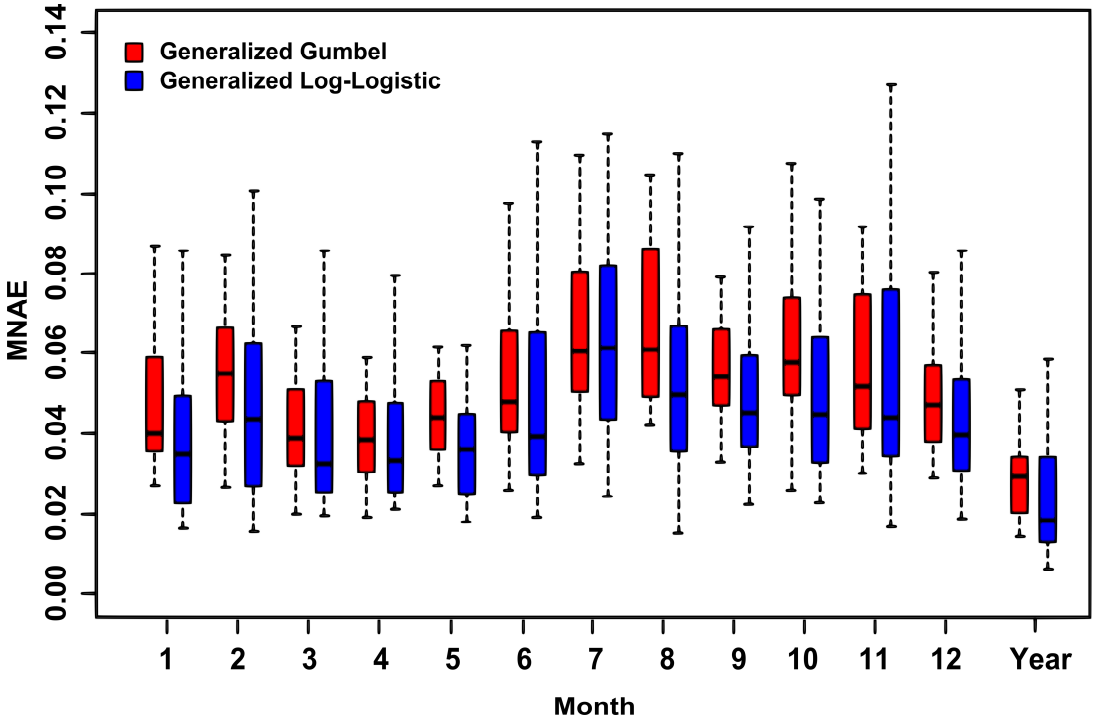

Figure 2. Comparison of Mean Normalized Absolute Error (MNAE) of the two probabilistic models used for the standardisation of the 45 stations with at least 40 years of data.

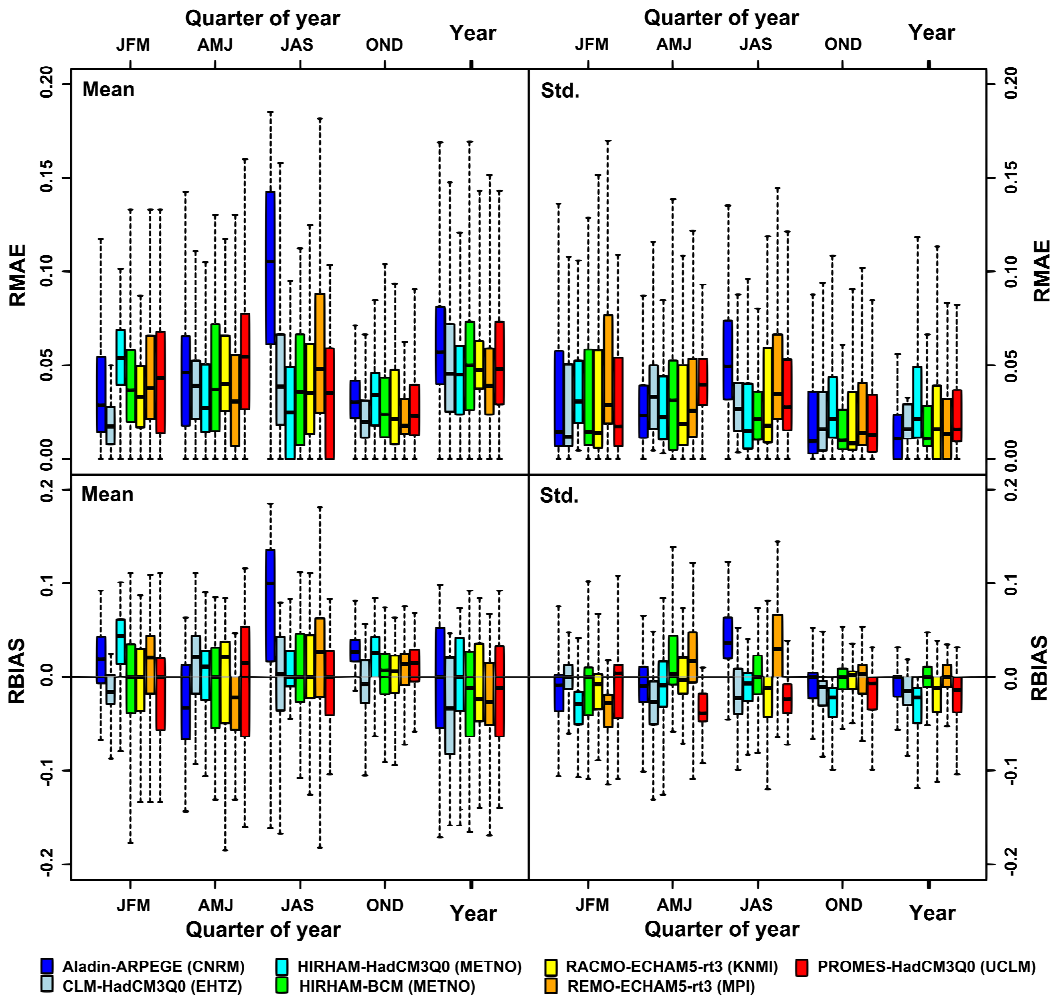

Figure 3. Relative mean absolute error (RMAE) and relative bias (RBIAS) of the average and standard deviations of simulated time series. Each bar represents the set of all stations according to each of the seven RCMs.

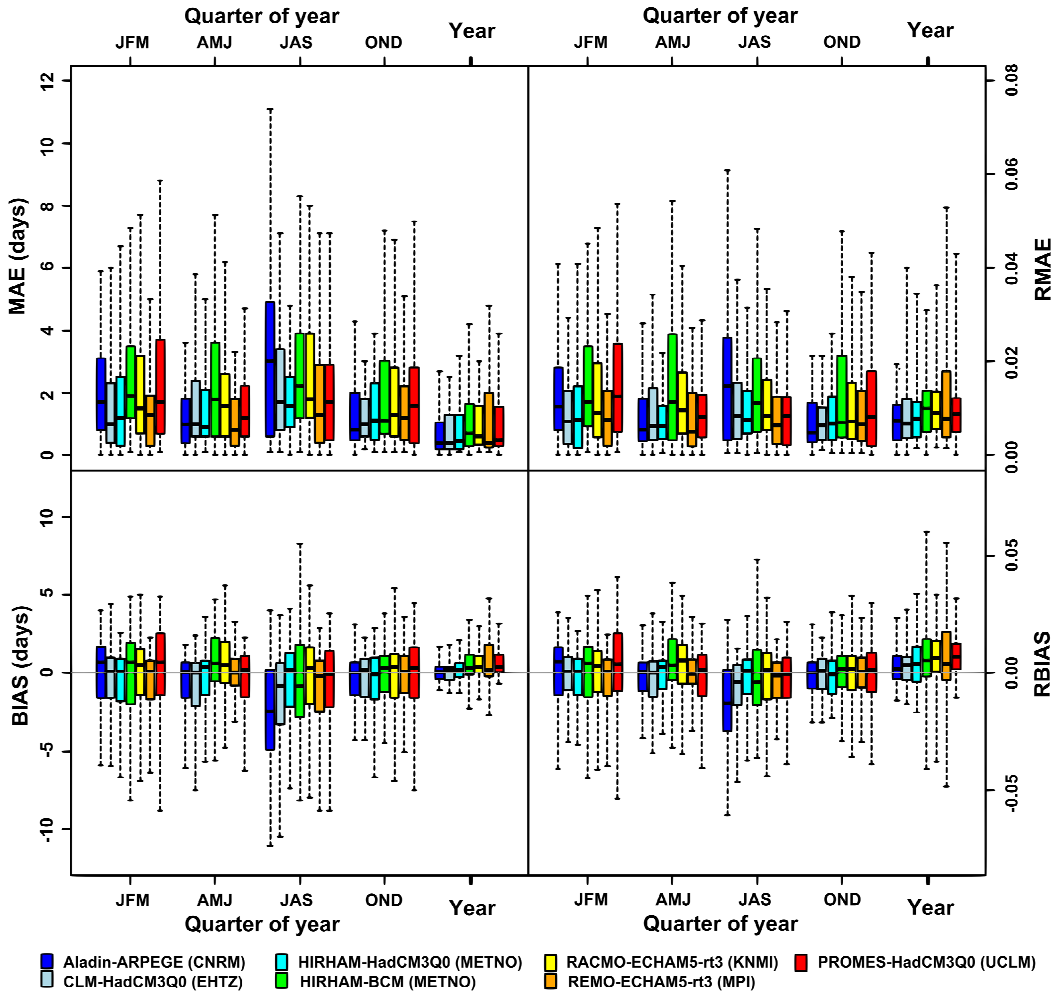

Figure 4. Absolute and relative mean error (MAE) and bias (BIAS) of the days without precipitation of the simulated series. Each bar represents the set of all stations according to each of the seven RCM.

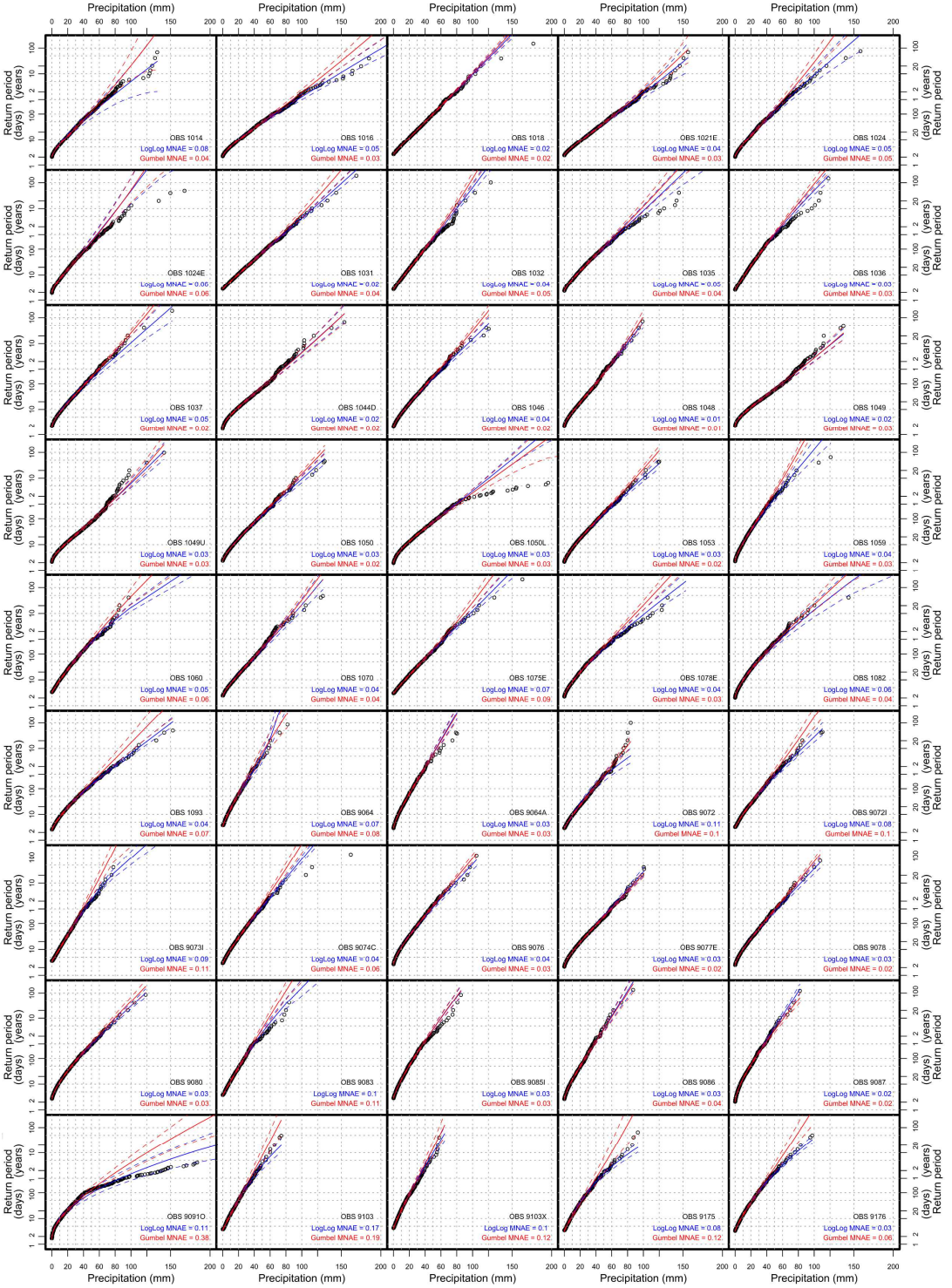

Figure 5. Return level plots of the daily precipitation for the used stations, in according to the two fitted models: modified Gumbel (red) and modified LogLogistic (blue). Dashed curves correspond to 95% confidence intervals; points correspond to empirical estimates excluding the maximum value of each time-series.

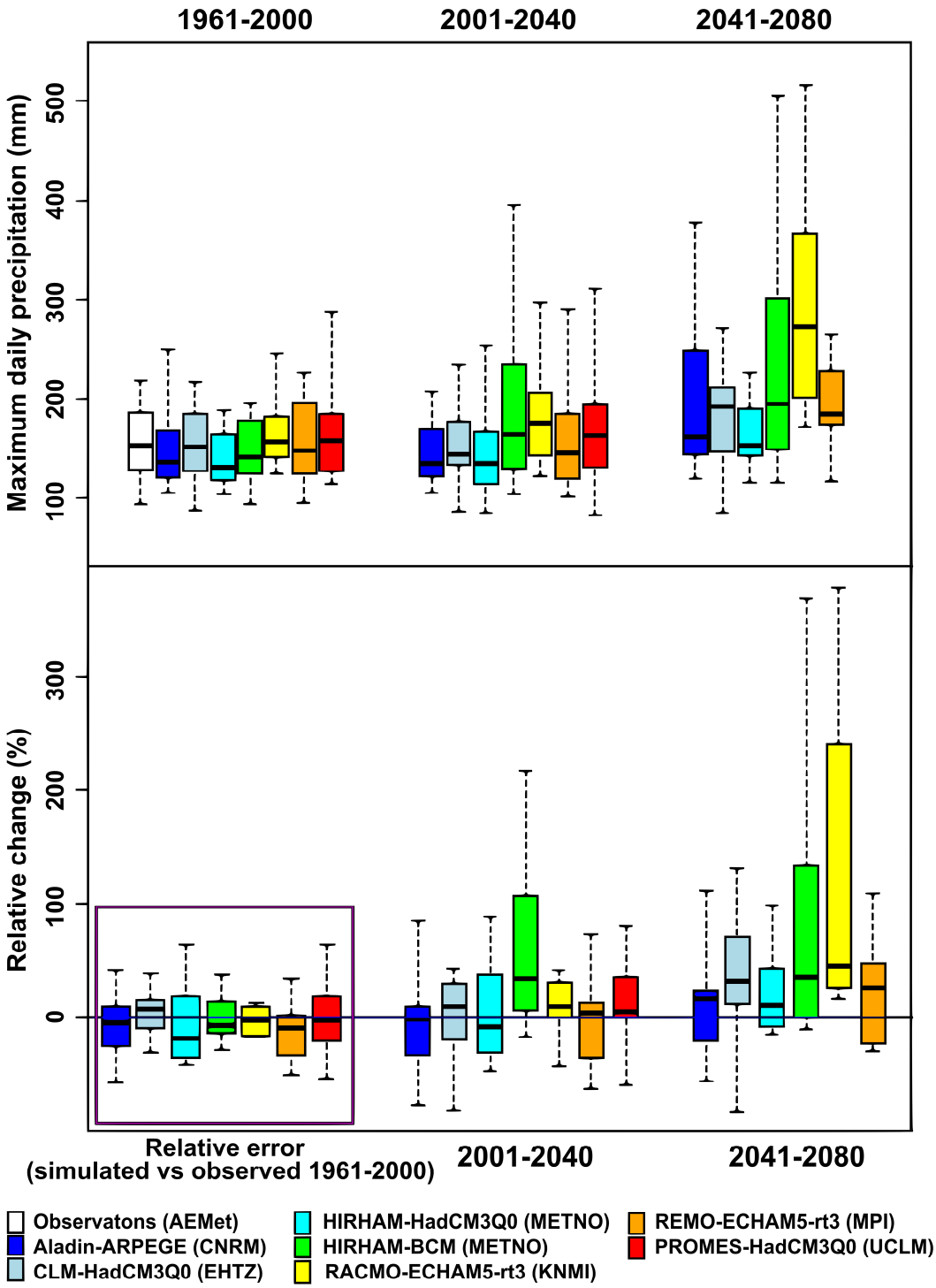

Figure 6. Projection of the expected daily rainfall for a return period of 100 years according to the theoretical curves fitted to the simulated stations. Above shows the projection of the absolute value for the three periods (1961-2000, 2001-2040 and 2041-2080). Below shows the relative change for 2001-2040 and 2041-2080 and the error of the simulation compared with observed time series (1961-2000).

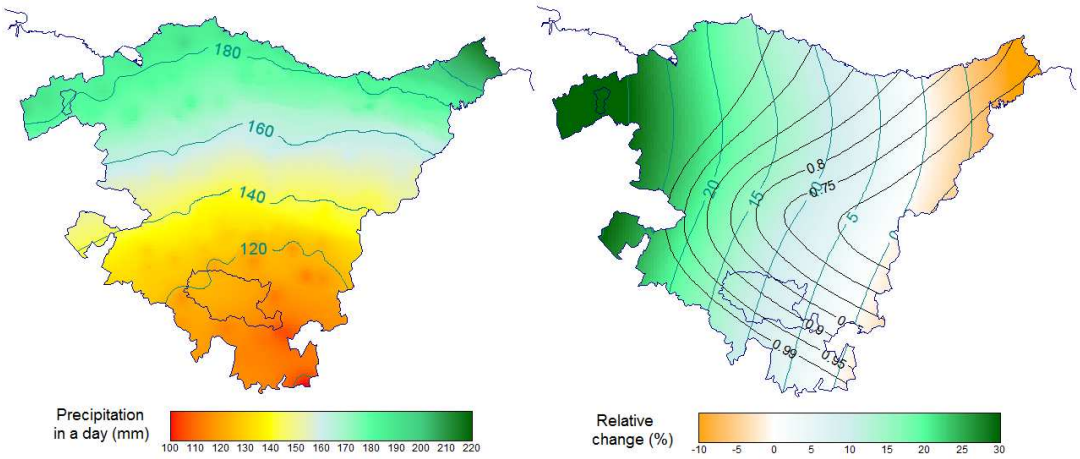

Figure 7. a) Maximum expected rainfall in one day for 1961-2000 simulated with a return period of 100 years. The values are estimated according to the theoretical curves fitted to the simulated stations, and they are interpolated by using a multiple spatial fit. b) Relative change of precipitation described in a), for the period 2001-2040 versus 1961-2000. The black lines represent the confidence intervals estimated from the statistics of the seven RCMs.

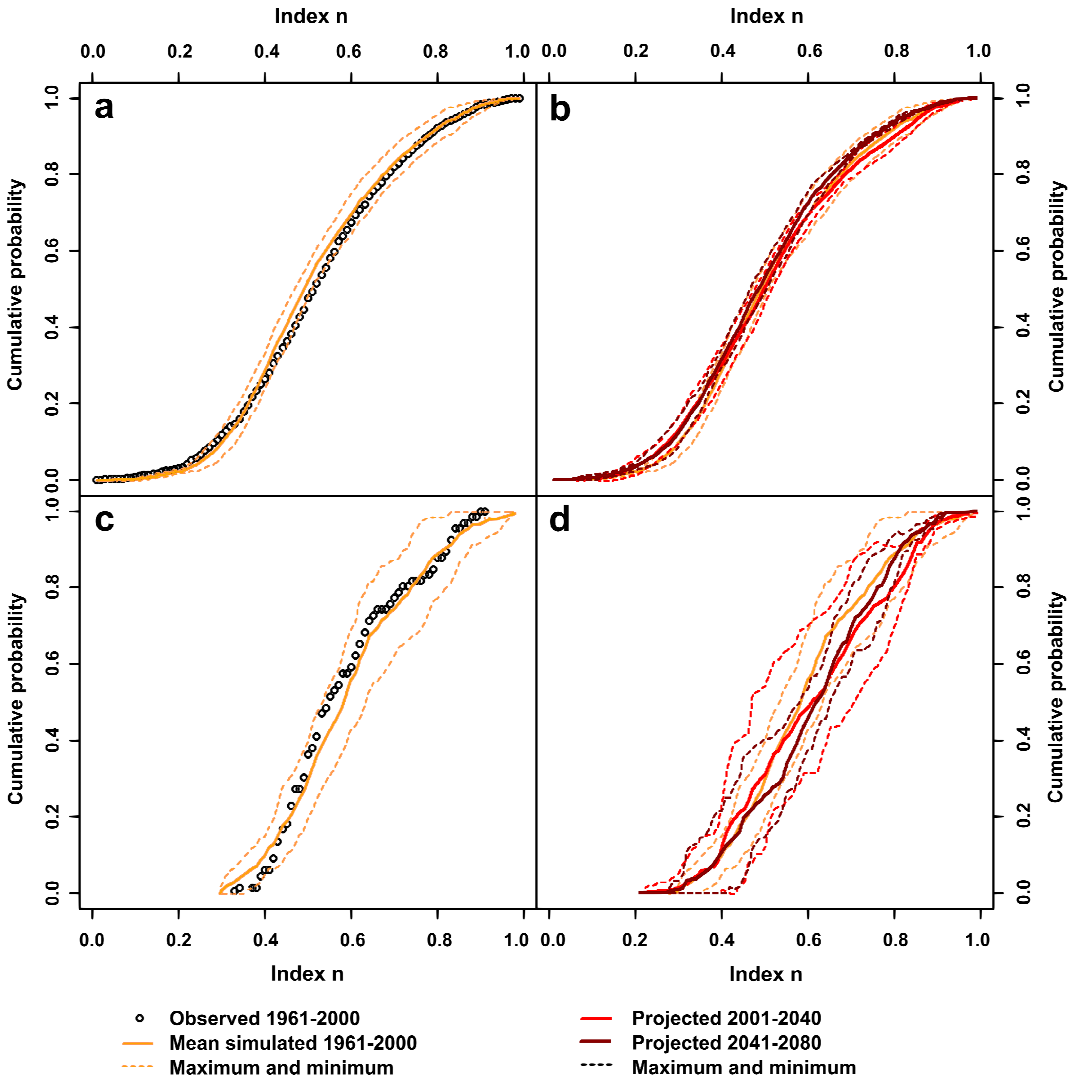

Figure 8. a) Empirical probability distribution of the index *n* to the set of all the stations in the period 1961-2000, and the range corresponding to the seven RCM simulations. b) Projection of change for 2001-2040 and 2041-2080. c) and d) Respectively, the same as a) and b) but only the rainfall exceeded 50 mm in one day.

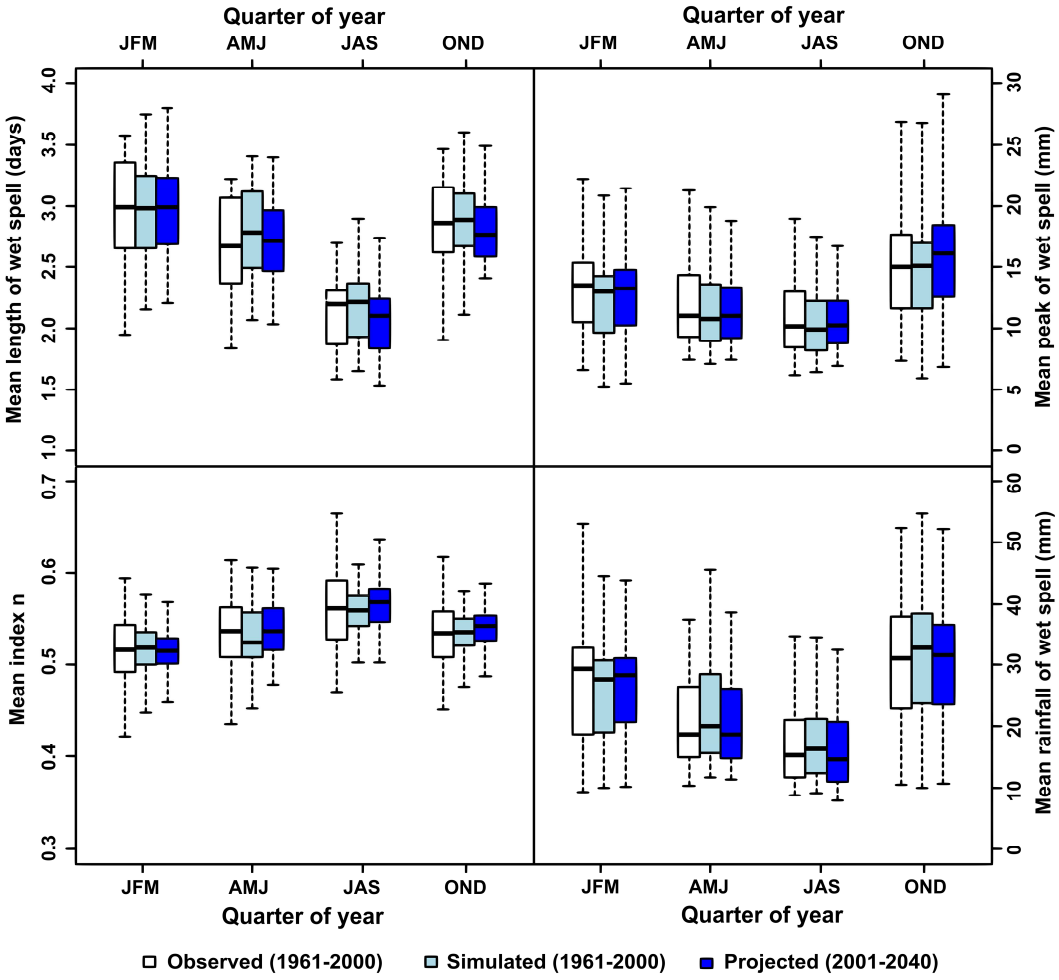

Figure 9. Comparison of observed and simulated wet spells for 1961-2000 and projected for 2001-2040. Each point of the boxplot represents the climatic value of a real or simulated station. Top left shows the average length of spells (days), top right, it shows the average daily peak (in millimeters). Bottom right shows the average total rainfall (in millimeters) and lower left compares the mean index *n* of wet spell.

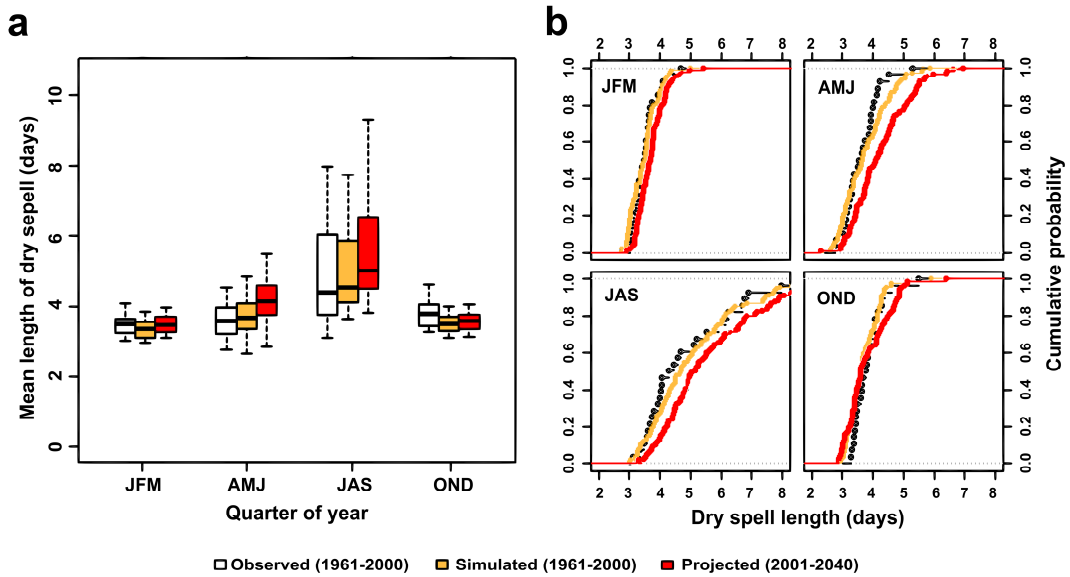

Figure 10. Evolution of dry spells for the period of 2001-2040 versus the period 1961-2000, both observed and simulated: a) boxplot of the average length of dry spells, and b) cumulative probability of the length of dry spells.