PERFORMANCE DE MODELOS CLIMÁTICOS REGIONAIS NA SIMULAÇÃO DE EXTREMOS DE PRECIPITAÇÃO E TEMPERATURA NA PENÍNSULA IBÉRICA

PERFORMANCE OF REGIONAL CLIMATE MODELS SIMULATING PRECIPITATION AND TEMPERATURE EXTREMES IN THE IBERIAN PENINSULA

Maria João Carvalho(1,2), Paulo de Melo-Gonçalves(1), Alfredo Rocha(1)
(1) Dept. de Física & CESAM – Centro de Estudos do Ambiente e do Mar, Universidade de Aveiro – Campus de Santiago, 3810-193 Aveiro.
(2) mariajcarvalho@ua.pt

SUMMARY

This work assesses the performance of the Regional Climate Models used in the European FP6 project ENSEMBLES, in simulating historical (1961-2000) annual and seasonal extremes of daily-total precipitation, daily-minimum and daily-maximum surface temperatures in the Iberian Peninsula, at an horizontal spatial resolution of approximately 25km. Two ensembles of simulations performed by these regional models are analysed in conjunction with the observed gridded dataset provided by the European Climate Assessment & Dataset. In one ensemble, all models downscaled ERA40 data, while in the second ensemble, each regional model downscaled at least one Global Climate Model simulation. Annual and seasonal statistics of daily extremes of precipitation and surface temperatures are quantified by the indices proposed by the CCI/CLIVAR/JCOMM Expert Team on Climate Change Detection and Indices. The performance of the models is accessed for both the mean state and interannual variability of these annual and seasonal indices. In addition, we also provide estimates of the recent-past climate change and the ability of the RCMs to reproduce it, taking into account the associated uncertainties.

1. Introduction

Climate Change (CC) plays an important role in several areas of everyday life (Field et al. 2012) and should therefore be studied in depth, both to understand the changes that the Planet has undertaken and those yet to come. While the recent-past can easily be studied by using observations, future climate changes can only be estimated through the use of Climate Models. Recently, the downscaling of Global Climate Models (GCM) has been done by using these to force Regional Climate Models (RCM) of higher resolution. In a study of models behaviour in the Alpine region, Prömmel et al. 2010 found that the use of a RCM had added value for 2-m temperature in the region, especially in regions of complex orography and coastal regions.

According to Déqué et al. (2007, 2012) there are four sources of error in climate simulations - sampling error, model uncertainty, radiative uncertainty and boundary uncertainty – of which the boundary forcing was found to be the most dominant for temperature data from the PRUDENCE project. They also added that the southern western Europe was the most affected area by the boundary uncertainty, whereas continental areas were more sensitive to the choice of RCM. A way to mitigate the uncertainty associated to the choice of RCM and GCM is to perform an ensemble of several RCM-GCM combinations (Gallardo et al. 2012). Furthermore, Maraun (2012) found that performing a BIAS correction or even eliminating the outlier simulations would improve the performance of the simulations.

The performance assessment of a model simulation cannot be assessed using only one climate statistical characteristic, since some are easier to reproduce than others. Maxino et al. (2008) used precipitation, maximum and minimum temperatures in order to evaluate the performance of the models submitted for the International Panel on Climate Change's (IPCC) Fourth Assessment (AR4) over the Murray-Darling basin in Australia. They attempted to do so by using the mean and Probability Distribution Functions (PDFs) of each of the variables and found that overall models would accurately reproduce the observed means but showed poorer results for the PDFs. Still using the AR4 maximum and minimum temperatures for the Australian region, Perkins et al. (2013) applied 3 different methods to assess differences in the 20 year temperature extremes as a function of model skill. The validation was performed through:

1. Difference between annual means;
2. PDF overlap;
3. Difference of the tail of the PDFs.

They found that a PDF or tail-based measure is preferable to the mean.

The study of climate change has increasingly focused on the extreme events of temperature and precipitation since these are more commonly cause for destruction and loss of life - for example the 2003 European heat wave (Bucker, 2005) and the Madeira Island extreme precipitation event (Luna et al., 2011).

2. Methodology

In order to compare the model results with observations, the simulations were split into two groups - ERA40-driven and GCM-driven. For each of the groups, an equitable weight ensemble was determined, along with the uncertainty using the Interquartile Range (IQR).

In order to achieve the goal of this work of assessing the performance of the simulations, two approaches were followed. Firstly, the seasonal climatologies of the variables were used to ascertain differences between observations and ensembles. Afterwards, extreme indices from the CCI/CLIVAR/JCOMM Expert Team on Climate Change Detection and Indices (ETCCDI) were used. This multi-perspective approach allows for a more comprehensive understanding of the simulations' behaviour since it studies, not only the mean but also the extremes of the variables. The indices used were the following:

- Consecutive Dry Days (CDD) – Greatest number of consecutive days with precipitation under 1 mm.
- Consecutive Wet Days (CWD) – Greatest number of consecutive days with precipitation over 1 mm.
- Extreme Temperature Range (ETR) – Maximum difference between maximum and minimum daily temperatures.

Furthermore, other extreme indices were considered:

- pr90p – number of days with precipitation above the 90th percentile of precipitation.
- tasmax90p – number of days with maximum temperature over the 90th percentile of maximum temperature.
- Tasmin10p – number of days with minimum temperature below the 10th percentile of minimum temperature.

The reference percentiles were determined for a control period between 1960 and 1990.

These indices were determined for each of the simulations and the observations, after which two equitable weight ensembles were determined: the ERA40-driven and the GCM-driven.

Taking the ensembles of these indices, their trend was determined using a simple linear regression, and compared to the observed trend. These trends had their statistical significance tested.

In order to assess the closeness of the ensembles' distributions to the observed ones, the Kolmogorov-Smirnov (K-S) test was applied to the original variables (precipitation, maximum and minimum temperature) as well as the ensembles of the indices. For the points where the ensemble distribution is considered to be statistically similar to the observed distribution at the 5% level, the test statistics was represented (Equation 1).

$$D_n = \max | M_i - O_i | \qquad (1)$$

Where M_i and O_i are the values of the ensemble mean and observed PDFs, respectively, at each of the bin values ($_i$).

3. Data

The European FP6 project ENSEMBLES (http://www.ensembles-eu.org/) provides a 16 member ensemble of climate simulations which can be divided into two sets: ERA40-driven (Regional Models driven by ERA40 reanalysis) and GCM-driven (Regional Models driven by Global Climate Models). Of all the simulations made available by the project, only those with the same grid were chosen, in order to avoid interpolation, which could act as an additional source of error.

RCM		Driving Global Climate Model					
		BCM	ECHAM5	HadCM			ERA 40
				Low S. 3Q3	Normal S. 3Q0	High S. 3Q16	
RACMO			X				X
Had RM	Low S.			X			
	Normal S.				X		
	High S.					X	
REMO			X				
RCA		X	X				

Table 1: Combinations of RCM and GCM used from those available from the ENSEMBLES Project

Also, observed data, E-OBS (V5.0) provided by the European Climate Assessment & Dataset (ECA&D) (http://eca.knmi.nl) was used. Recent-past (1961-2000) daily values of minimum and maximum temperature as well as total precipitation were used on a rotated grid of 0.22º x 0.22º. The integration domain used by the ENSEMBLES Project includes the entire European continental territory. However, the focus of this work was a domain containing only the Iberian Peninsula, whose topography can be seen in figure 1.

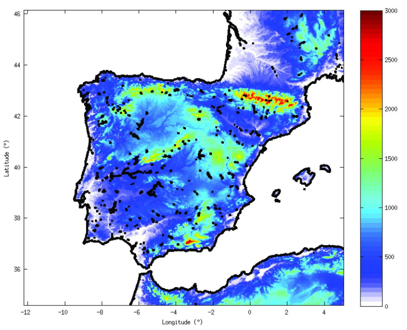


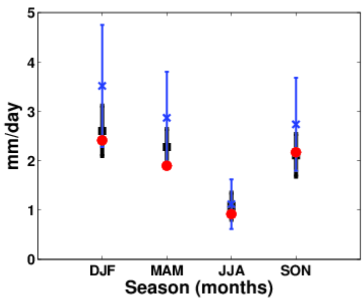
Figure 1: Topography of the domain (m) using the USGS GTOPO 30 database.

4. Results & Discussion

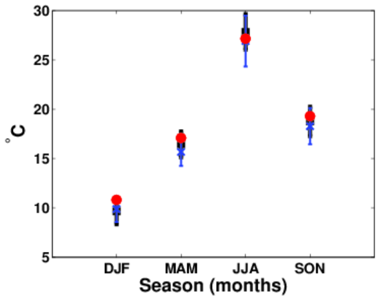
The first thing to look at is the ensemble and observed climatologies of the original variables: precipitation, minimum and maximum temperatures. By calculating the spatial mean of the seasonal climatologies these three sets of data together with the uncertainties of the ensembles, one can have a general idea of their relative behaviour. As can be seen in figure 2, for all three variables the ERA40-driven ensemble has better performance (simulations closer to observations) than the GCM-driven ensembles, as would be expected. Furthermore, GCM-driven ensemble shows far larger (approximately twice as large) uncertainty when compared to ERA40-driven ensemble. These differences are especially evident for precipitation while as for temperature, GCM-driven and ERA40-driven ensembles show closer results when comparing to observations. The only exception is summer precipitation which shows lower uncertainty and better performance by both ensembles, which is easily explained by the fact that summer is the driest season in the domain.

Although this work was developed for both ERA40-driven and GCM-driven ensemble means, due to space constraints only results for the latter will be presented hereafter although a short discussion of both will be introduced in the conclusion and therefore, the mention of ensemble will be referring to the GCM-driven ensemble.

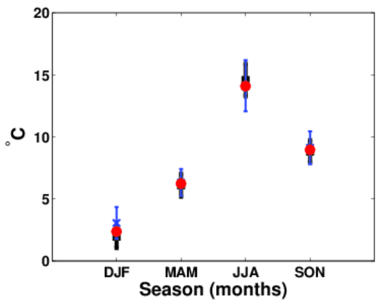
The analysis of the spatial mean of seasonal climatology of the observed and GCM-driven indices showed very different results from those obtained for the original variables. While the maximum and minimum temperatures ensemble showed better performance than the precipitation one, the opposite happens to the indices.



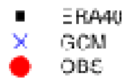
a) *Precipitation*



b) *Maximum Temperature*



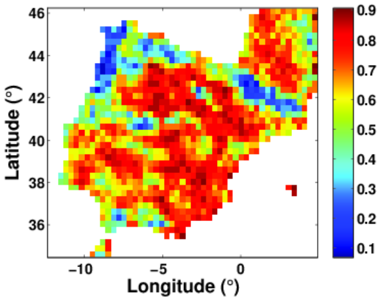
c) Minimum Temperature



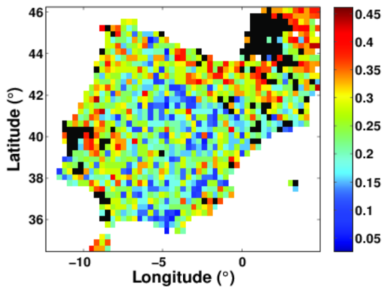
<u>Figure 2</u>: Spatial Mean of the climatology of the variables using the GCM-driven and ERA40-driven ensembles (and their uncertainty) as well as observations.

Both the number of consecutive dry and wet days (CDD and CWD) proved to be well represented by the ensemble, since the ensemble mean values of the climatologies' spatial mean were close to the observed ones (and inside the uncertainty of the ensemble). Even so, the uncertainty surrounding the ensemble calculation was high, neighbouring 30%. On the other hand, pr90p showed far worse results, with the consistent gross underestimation of the index by the ensemble, even though the latter presented with low uncertainty. This result points to good ensemble capability of estimating the ratio of rain/no-rain days with difficulties in the determination of rain amount.

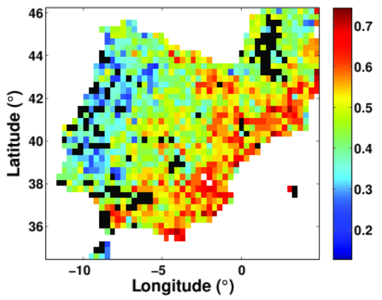
Temperature indices proved to be less accurate than CDD and CWD. While tasmax90p and tasmin10p are grossly overestimated by the ensemble (by approximately 4 º C) ETR is underestimated by ~6 days. Even though the ensemble uncertainty for these indices is lower (on average 11% for both tasmin10p and tasmax90p and 20% for ETR) observation points were without exception, outside the uncertainty bars. Therefore, all models produce more warmer (tasmax90p) and colder (tasmin10p) days, while simulating lower overall difference between maximum and minimum temperature.



a) *pr90p*



b) *CDD*



c) CWD

Figure 3: K-S test statistics (D_n) between the ensemble mean GCM-driven and OBS, using the distributions of yearly values of each of the indices: a) pr90p, b) CDD and c) CWD. For black grid boxes, differences between PDFs are not statistically significant (5%) and D_n is not shown.

As happened for the spatial mean of climatologies, there are clear differences in the ensemble performance for precipitation and temperature indices when their PDFs are analysed. As mentioned in the Methodology section, the K-S test was applied in order to assess the grid points where each of the indices' modelled distribution was considered as statistically similar to the observed one at the 5% confidence level.

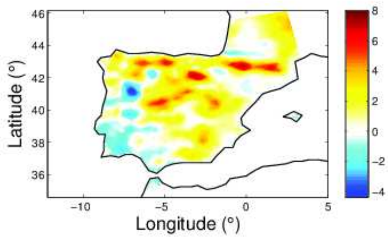
Precipitation indices show that the index modelled PDF is similar to the observed one at the 5% confidence level in most of the grid points (all of them for pr90p and most of them for the other two indices) – figure 3. For the points where the null hypothesis (distributions are the same) cannot be rejected, the test statistics (D_n) was represented. For pr90p, the lowest D_n values are found to be at the northern part of the Iberian west coast and along the Pyrenees. On the other hand, the lowest K-S test statistics for CWD are found mostly in Portuguese territory while Spain presents higher (mostly along the eastern coastline) differences between modelled and observed PDF. The K-S test statistics field for the number of dry days (CDD) does not show a clear pattern of lower and higher values, albeit having the overall lowest (approximately half) values when compared to the two other indices.

When comparing the temperature indices' modelled and observed PDFs using the K-S test (not shown), results show worst ensemble performance. For both tasmax90p and tasmin10p, most of the grid points reject the null hypothesis "the distributions are the same" and the remaining ones show no clear pattern of lower/higher D_n values. For ETR, on the other hand, although also having a high percentage of points with statistically different distributions, shows that higher values of the test statistics can be found along the southern part of the west Portuguese coastline, with the rest of the domain presenting lower, close to zero D_n.

In order to further understand the differences between the GCM-driven ensemble and observations, the seasonal climatologies of the indices were calculated for each of the two sets of data and the difference field (GCM - OBS) was represented.

The pr90p and CDD difference fields (not shown) show that the underestimation of these indices by the ensemble is consistent in the domain, with only two clear exceptions:

- Pr90p summer difference field has a horizontal band along the southern coastline and that extends for about 300 km north of lower absolute values.
- CDD spring difference field shows lower difference values (more negative) in the north/north-west region of the IP.



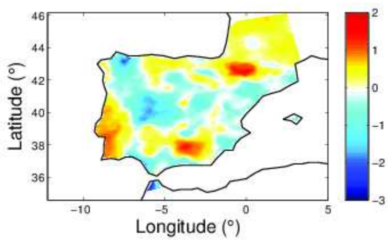
<u>Figure 4</u>: CWD difference field between GCM-driven ensemble and observations for the winter.

CWD difference fields (for winter these fields are shown in figure 4) show that the overestimation of these indices is larger in areas of complex topography, especially in the Pyrenees. Furthermore, for summer and autumn, the difference fields also present a clear west/east pattern of negative/positive values.

As for the temperatures indices' difference fields, results become more varied (not shown). For tasmax90p, while winter and spring show almost exclusively negative values (except a small area in the Pyrenees during winter), summer shows positive values along the western and northern areas and autumn shows a spatially consistent overestimation, with the same magnitude of the underestimation during the other seasons.

The difference fields for tasmin10p seasonal climatologies, show more seasonal and spatial variations. Spring is the only season which shows the same signal (positive) in the entire domain, with higher values in the south and north-east. These regions are the ones showing positive difference values in the summer, in an otherwise positive field.

The winter difference field can be seen in figure 5. As can be seen, there are three areas of high positive differences: south western coastal region, the Sierra Nevada and the Pyrenees. The summer field is similar to the winter but with lower differences.



<u>Figure 5</u>: Tasmin10p difference field between GCM-driven ensemble and observations for the winter (DJF).

Extreme temperature range (ETR) shows a clear dichotomy of mostly positive/negative values for summer and autumn/winter and spring.

Other skill measures such as the Root Mean Square Error (RMSE), the Standard Deviation of the ERROR (STDE) and Pearson Correlation were applied both to the ensemble, as well as the individual simulations. However, due to space constraints, those results are not presented here.

5. Conclusion

Climate simulations are an important tool in order to foresee changes that will occur in the future, especially with the impact that these can have on everyday life issues such as water and energy management, agriculture, fisheries and health. However helpful, there is high uncertainty associated with the results, be it from model uncertainty or the radiative forcing. Therefore, it becomes important to assess the ability of the models to reproduce the current and recent-past climate in order to understand their shortfalls.

By comparing model results with an observational database, there is the possibility to assess the variables and regions where models are less accurate and therefore, where their results should be taken into account more carefully. Since the extreme events are a major source of destruction and loss of life, this work focused on them.

When using simulations from GCMs to force RCMs, there are already two sources of error. As this analysis showed, there is added shift from of the GCM-driven ensemble from observations, when compared to the ERA40-driven ensemble. Furthermore, the uncertainty of the first is also far greater than that of the latter.

From the comparison between precipitation and temperature, an interesting conclusion can be drawn. Although there is better performance by the maximum and minimum temperatures ensembles, precipitation indices outperform the temperature ones.

Lastly, the ensembles showed lower performance in areas where the topography is complex (high mountainous regions such as the Pyrenees) and along the coastlines. These may be due to the resolution of the lower boundary condition and the water-land frontier.

Acknowledgments

This study was supported by FEDER funds through the Programa Operacional Factores de Competitividade – COMPETE and by Portuguese national funds through FCT – Fundação para a Ciência e a Tecnologia, within the framework of Projects “Climate change of precipitation extreme episodes in the Iberian Peninsula and its forcing mechanisms” – CLIPE. PTDC/ AAC-CLI/111733/2009 and “Urban Atmospheric Quality, Climate Change and Resilience.” EXCL/AAG-MAA/0383/2012

References

Bucker, G. 2005. Vulnerable populations: lessons learnt from the summer 2003 heat wave in Europe. *Europe Surveillance*, 10(7):147.

Déqué, M., Rowell, D., Lüthi, D., Giorgi, F., Christensen, J., Rockel, B., Jacob, D:, Kjellström, E., de Castro, M. and van den hurk, B. 2007. An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections. *Climate Change,* 81: 53-70.

Déqué, M., Somot, S., Sanchez-Gomez, E., Goodess, C., Jacob, D., Lenderink, G. and Christensen, O. 2012. The spread amongst ensembles regional scenarios: Regional Climate Models, driving general circulation models and interannual variability. *Climate Dynamics*. 38(5): 951-964.

Field, C., Barros, V., Stocker, T., Qin, D., Dokken, D., Ebi, K., Mastrandrea, M., Mach, K., Plattner, G. and Allen, S. 2012. Managing the risks of extreme events and disasters ti advance climate change adaptation. A special report of Working Groups I and II of the Intergovernmental Panel on Climate Change. *Cambridge University Press, Cambridge, UK.*

Gallardo, C., Gil, V., Hagel, E., Tajeda, C. and de Castro, M. 2012. Assessment of climate change in Europe from an ensemble of Regional Climate Models by the use of Köppen-Trewartha classification. *International Journal of Climatology*.

Luna, T., Rocha, A., Carvalho, A., Ferreira, J. and Sousa, J. 2011. Modelling the extreme precipitation event over Madeira Island on 20 February 2010. *Nat. Hazards Earth Syst. Sci.*, 11: 2437-2442.

Maraun, D. 2012. Nonstationarities of Regional Climate Model biases in european seasonal mean temperature and precipitation sums. *Geophysical Research Letters*, 39(6): L06706.

Maxino, C., McAvaney, B., Pitman, A. and Perkins, S. 2008. Ranking the AR4 climate models over the Murray-Darling basin using simulated maximum temperature, minimum temperature and precipitation. *International Journal of Climatology*, 28(8): 1097, 1112.

Perkins, S., Pitman, A. and Sisson, S. 2013 Systematic differences in future 20 year temperature extremes in AR4 model projections over Australia as a function of model skill. *International Journal of Climatology*, 33: 1153-1167.

Prömmel, K., Geyer, B., Jones, J. and Widmann, M. 2010. Evaluation of the skill and added value of reanalysis driven regional simulation for alpine temperature. *International Journal of Climatology*, 30(5): 760-773.