Collaborative Project

CLIM-RUN

Climate Local Information in the Mediterranean region Responding to User Needs

WP 3 – Observational support and downscaling methods
Task 3.4 Downscaling methods and portal

Adaptation of the ENSEMBLES downscaling portal to CLIM-RUN requirements

Project No. 265192– CLIM-RUN

Start date of project: 1st March 2011 Duration: 36 months

Organization name of lead contractor for this deliverable: University of Cantabria

Due Date of Deliverable: 1st March 2012 Actual Submission Date: 31st March 2012

Authors: M. D. Frías, A. Cofiño, A. Casanueva and C. Sordo

Table of Contents

1. Introduction

Global climate models (GCMs) are able to reasonably simulate main global characteristics, however they are not as accurate when it comes to the simulation of more local details requested in many impact studies (agriculture, fire, energy, tourism or health). Despite an increasing ability to successfully model present-day climate, the latest generation of GCMs still has serious difficulties in capturing regional variability details in smaller regions. The GCM restrictions in regional scales can be partly explained by their limited horizontal resolution (typically 250 Km), model uncertainties and the crude representation of the orography (Wigley et al, 1990; Storch et al, 1993). This make it necessary to use specific regionalisation techniques or downscaling to bridge the gap between the coarse resolution of the global models and the high resolution required by end--user applications. Statistical downscaling is a sound and mature field which allow us to derive regional or local scale information (typically surface variables such as precipitation or temperature) from integrations of coarse resolution from the GCMs (typically large scale fields such as 500 mb geopotential height).

Statistical downscaling is nowadays a complex multidisciplinary discipline involving a cascade of different scientific applications to access and process large amounts of heterogeneous data. Therefore, interactive user-friendly tools are necessary in order to ease the downscaling process for end users, thus maximizing the exploitation of the available climate projections. The ENSEMBLES downscaling portal was initially developed within the EU-funded ENSEMBLES project (2004-2009) following an end-to-end approach. Afterwards, a complete reimplementation (version 2) was performed to ensure the appropriate adaptation of the portal (different views for different users). This document describes the adaptation of this portal made for CLIM-RUN according to the case studies involved in the project: tourism, energy and wild fire. To this end, as described in detail bellow, particular datasets required from the CLIM-RUN WPs related to those case studies are now available in the portal. There is also a user guide for end-users with detailed information on the steps to be followed to undertake a particular downscaling experiment using the downscaling portal (http://www.meteo.unican.es/downscaling/climrun). This user guide is intended for end-users with some basic knowledge on statistical downscaling and focus on the steps to be followed to undertake a particular downscaling experiment using the downscaling portal (Gutiérrez et al 2011). We soundly recommend reading this document to all the CLIM-RUN partners interested in applying statistical downscaling using the portal. The present deliverable complements this user guide providing further information about the validation process produced by the portal.

2. Adaptation of the portal to CLIM-RUN

The ENSEMBLES downscaling portal has been adapted to CLIM-RUN taking into account the requirements made by the CLIM-RUN partners involved in the case studies. Any participant in this project can access to this portal through the link **http://www.meteo.unican.es/downscaling/climrun**. Figure 1 shows the web of access to the portal. First, any user is required to be registered as member of the project selecting the option marked as "*Register in the portal*". A new page is loaded to fill different profile items such as the

project partner, full name, the username desired, password, etc as shown in Figure 2. After this previous step, a confirmation email will be sent allowing the user to access to the downscaling portal as CLIM-RUN partner.

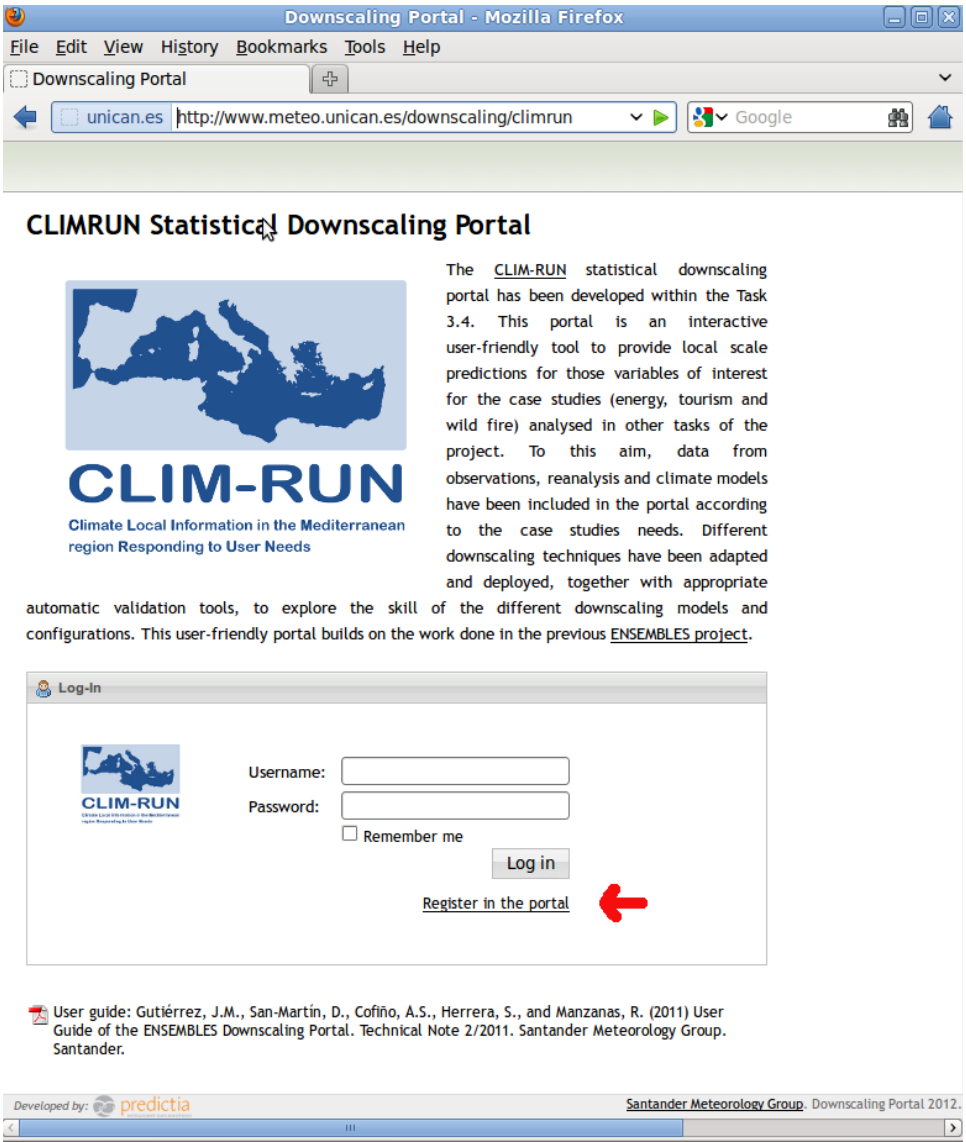

Figure 1: Web of access to the portal

After login to the portal a first window ("My home") is shown and the user is ready to perform the downscaling (see Figure 3). As an illustrative example, the portal includes a "demo" experiment (Iberia demo) , which focuses on maximum temperature in five locations/cities for the 2091-2100 decade. This experiment is soundly described in the user guide of the portal and it is available for all users (in write-protect mode). It can be followed step to step through the different panels of the portal in order to see a typical application.

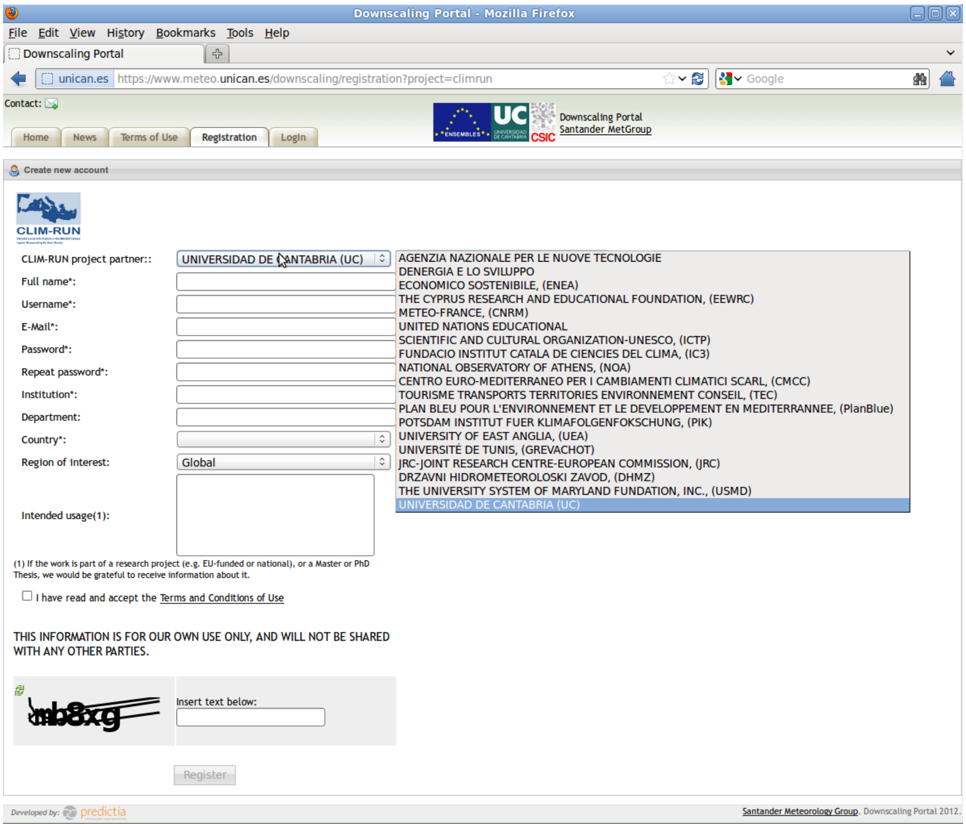

Figure 2: Registration in the portal

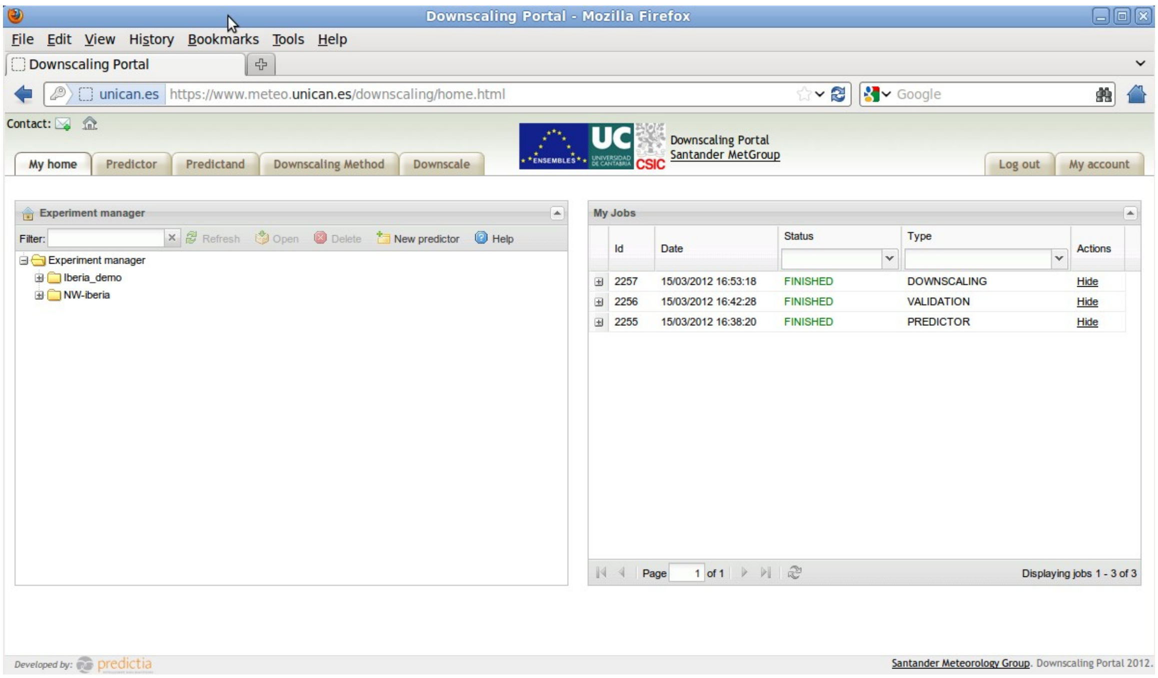

Figure 3: Main window of the downscaling portal

2.1.Data

Three are the case studies analysed in CLIM-RUN: tourism, energy and wild fire focused in different areas over the Mediterranean region. At this first stage, only partners from the wild fire case study have required particular climate variables to be available in the portal for downscaling purposes: instantaneous values of temperature, relative humidity and wind velocity and accumulated precipitation – all for noon. These values are related to the estimation of the Canadian forest Fire Weather Index (FWI) (Stocks et al, 1989) which they need at local scale. In particular, the areas of interest for wild fire studies in CLIMRUN are Greece and Spain. According to this request, the most suitable and available datasets have been uploaded in the portal: ERA-Interim and WATCH Forcing Data. These datasets are available in the CLIM-RUN portal for the whole Europe since some of these variables could also be interesting for the other two case studies applied in other Mediterranean regions (Savoie-France, Tunisia and Cyprus for tourism and Spain, Morocco and Cyprus for energy). It is expected that partners working on energy and tourism case studies will start to make their data requests in the next months. The downscaling portal is now ready to upload any other variable or dataset according to following requests. As included in the user guide, users can also upload their own data for particular locations. These data will be only available for their personal use, unless the user gives open access to the data.

The Water and Global Change (WATCH) Forcing Data (WFD) is a dataset generated in the framework of the EU WATCH project (www.eu-watch.org, 2007-2011). The aim of this project was the analysis, quantification and prediction of the components of the current and future global water cycles and related water resources states, the evaluation of their uncertainties and the clarification of the overall vulnerability of global water resources related to the main societal and economic sectors. WATCH results provide global sub-daily meteorological data over land for the whole of the twentieth century at 50km grid scale resolution and for the future (21^{st} century). Surface variables from the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) product were considered as the basis data used in the derivation of this dataset for the period 1958 to 2001 and from reordered ERA-40 data for the period 1901 to 1957. The variables included in this dataset are wind speed at 10 m, air temperature at 2 m, surface pressure, specific humidity at 2 m, downward longwave radiation flux, downward shortwave radiation flux, rainfall rate, and snowfall rate. The variables are stored at 3 or 6 hourly time steps depending on the case. More information about this data can be found in Weedon et al (2011). Only noon values of instantaneous temperature, relative humidity and wind velocity and accumulated precipitation were selected in the CLIM-RUN downscaling portal from the period 1961 to 2000 for the whole Europe. Values for relative humidity were calculated using the formulation from Wallace and Hobbs (1977) using the specific humidity. As mentioned above, these four variables are related to the FWI which is necessary at local scale for the wild fire case studies defined in CLIM-RUN. This dataset is named in the portal as WATCH-Fire to remark the particular characteristics of the climate variables included.

Particular data from ERA-Interim are also available in the downscaling portal. ERA-Interim is the latest global atmospheric reanalysis produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) conducted in part to replace ERA-40 reanalysis. Nowadays it covers the

period from 1979 to 2011, although it is pretended to extend back to the early part of the twentieth century. ERA-Interim provides a large variety of 3 hourly surface variables, among other variables, at a 1.5º resolution. This reanalysis is derived from successive short-term integrations of a general circulation model that assimilated different kind of data (satellite data and land and sea surface observations). More information about this product can be found in Dee et al (2011). As for the WFD dataset, only those variables related to the FWI are available in the portal at this stage. In this case, relative humidity was calculated using the formulation from Bohren and Albrecht (1998) taking into account, the dew point and the temperature. According to the particular characteristics of these four variables, this dataset is renamed in the portal as ERAInt_Fire. More information on this dataset and its suitability for fire danger analysis can be found in Bedia et al (2012).

Due to the spatial and temporal coverage of this dataset, more variables from ERA-Interim will probably be included in the portal in the next months for applications in other case studies.

As in the ENSEMBLES downscaling portal there are some other datasets from observations and climate models available for CLIM-RUN partners. GSN world dataset and the GSOD dataset from the National Climatic Data Center for total precipitation and mean, maximum and minimum temperature over Europe are also in the portal for the present period. Several predictors from three different climate models from AR4 (BCM2, CNCM3 and ECHAM5) and one model from ENSEMBLES (HADGEM2) can be used for downscaling purposes. In particular, data of temperature, U and V velocity, specific humidity, mean sea level pressure and geopotential at different levels are available for the 20C3M and A1B emission scenarios.

3. Validation

The downscaling portal user guide provides to the users a brief introduction of the validation process produced by the portal (Figure 4). The second aim of this deliverable is to increase this information providing a more detailed description of the validation process generated. It makes an attempt to help the user to properly analyse the statistical scores calculated for the validation of the downscaling method applied.

Two validations are performed in the portal, one on a daily basis and the other on a 10-day-aggregated one (using the mean as inference method to aggregate the data). Note that, depending on the user's needs, both time-scales might be useful. The corresponding statistics obtained for observations and predictions and those calculated for the quality of the forecasts (*accuracy* and *distributional similarity*) are described bellow. More information about these scores can be found in Jolliffe and Stephenson (2003) and Von Storch and Zwiers (2001).

3.1. Observations (Predictions) Statistics

- RR: Rainfall Rate. It measures the frequency of wet days and it is calculated as the number of wet days divided by the size of the sample, n, expressed in %.

$$RR = \frac{n_{wet}}{n}100$$

This score is specific for precipitation. The threshold considered for defining wet days is 0.1 mm.

- Mean: Arithmetic mean. It measures the central tendency in a sample. It is calculated as the sum of all data points (x_i, i=1,...n) divided by the size of the sample (n). The arithmetic mean is greatly influenced by outliers. For this reason, robust statistics such as the median may provide a better description of central tendency.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

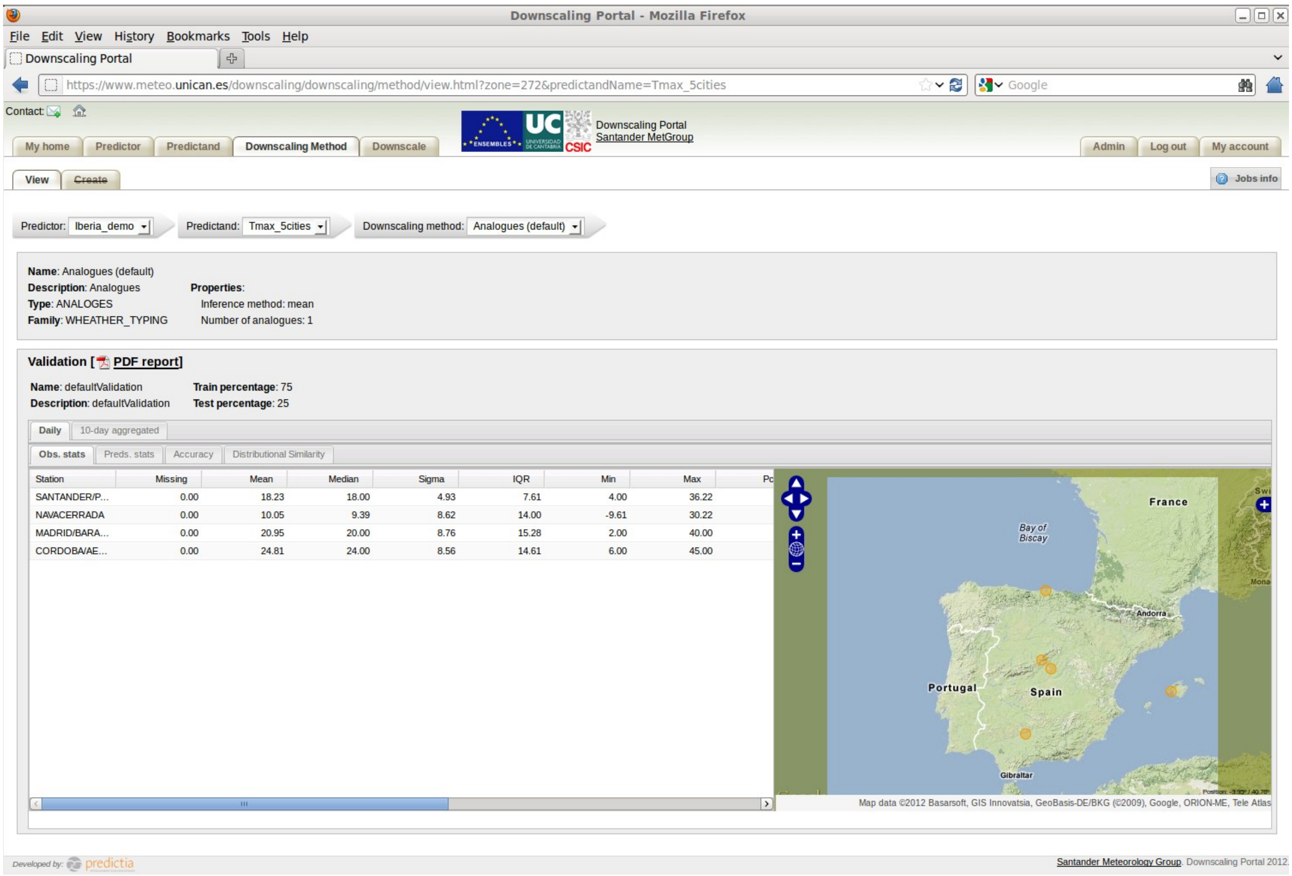

Figure 4: Window to access to the results from the validation of the downscaling method

- Median: Median measures the central tendency in a sample. It is described as the value separating the higher half of the sample from the lower one (50th percentile). It can be found by arranging all the values from the lowest to the highest and picking the middle one. For data symmetrically-distributed, the mean an the median are the same.

- Sigma: Standard Deviation (also denoted as Std). It measures how spread out the data are. Basically, a large standard deviation indicates that the data are far from the mean and a small standard deviation indicates that they are clustered closely around the mean. The

standard deviation is the square root of the variance. A useful property of the standard deviation is that, unlike variance, it is expressed in the same units as the data.

$$Std = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

- IQR: Interquartile range. It measures the spread of a sample being equal to the difference between the upper (75th percentile) and lower (25th percentile) quartiles, Q1 and Q3 respectively. The interquartile range of a sample can be calculated by integrating the probability density function (f). The lower quartile, Q1, is a number such as the integral of the probability density function (PDF) from -∞ to Q1 equals to 0.25,

$$\int_{-\infty}^{Q_1} f(x)\, dx = 0.25$$

while the upper quartile, Q3, is a number such as the integral from -∞ to Q3 equals to 0.75.

$$\int_{-\infty}^{Q_3} f(x)\, dx = 0.75$$

The interquartile range is used to build box-plots, simple graphical representations that shows with a box the spread of the data falling between the 25th and 75th percentiles.

- Min: Minimum. The smallest value in the data.

- Max: Maximum. The largest value in the data.

- PX: X^{th} percentile. Value below which X% of the data points are found. X=10, 90.

- Missing: Percentage of missing values within the data. Ranged in [0, 100]. Perfect score: 0.

3.2. Accuracy

Accuracy is one of the main aspects that must be examined when looking at the quality of a forecast. Essentially, it is the level of correspondence, i. e., punctual agreement, between forecasts and observations. Differences between both are errors. The lower the errors, the greater the accuracy. Note that some of the scores are presented in units of some descriptive statistic, which allows for direct comparison among stations and/or seasons, not worrying about their different regimes. In particular, those scores re-scaled by the Mean (Sigma) are named with a n (N) at the beginning of their names.

- HIT: Hit Rate. It is the probability of occurrences (i. e. wet day) that were correctly forecast. This score ranges in [0, 1] being 1 the perfect score.

$$HIR = P(For = 1|Obs = 1)$$

- FAR: False Alarm Rate. It is the probability of non-occurrences that were incorrectly forecast. This score ranges in [0, 1] being 0 the perfect score.

$$HIR = P(For = 1|Obs = 0)$$

  Note that both scores, HIR and FAR, are only calculated in the portal for the case of the daily precipitation. They are not evaluated for the 10-daily validation since aggregated data are considered to be continuous. HIR and FAR must be considered together in order to validate the discrete part precipitation. The threshold considered for defining wet days is 0.1 mm.

- rho: Pearson's Product-Moment Correlation Coefficient. It measures the strength of the linear relationship between observations and forecasts. The Pearson correlation coefficient between two variables (x and y, observations and forecasts in our case) is defined as the covariance of the two variables divided by the product of their standard deviations:

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

  The Pearson correlation coefficient measures how close the points of a scatter plot (observations against forecasts) are to a straight line. This score ranges in [-1, 1]. A value of 1 (-1) implies that a linear equation describes the relationship between observations and forecasts perfectly, with all the data points lying on a line for which forecasts increase (decrease) as observations increase. A value of 0 implies that there is no linear correlation between the variables. This score does not take bias into account, i. e., it is possible for a forecast with large errors to account for a good Pearson correlation coefficient respect to the observations. This score is sensitive to outliers.

- r: Spearman's Rank Correlation Coefficient. It measures the dependence (through some monotonic function) between observations and forecasts. The Spearman's Rank correlation coefficient is defined as the Pearson correlation coefficient considering the ranked variables.

  The sign of this score indicates the direction of association between observations and forecasts. A positive (negative) coefficient indicates that forecasts tend to increase (decrease) as observations increase. Its magnitude increases as observations and forecasts become closer to being perfect monotone functions of each other. A Spearman correlation coefficient of 1 (-1) results when observations and forecasts are monotonically related, even if their relationship is not linear. In contrast, this does not yield to a perfect Pearson correlation. The Spearman correlation coefficient is less sensitive than the Pearson correlation coefficient to outliers that may be in the tails of both observations and/or predictions. This score should be used when validating precipitation rather than the Pearson correlation coefficient.

- MAE: Mean Absolute Error. It measures the average magnitude of the forecast errors and ranges in [0, ∞) being 0 the perfect score. It does not indicate the direction of the deviations.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |For_i - Obs_i|$$

- nMAE: Mean Absolute Error (MAE), in units of the observed mean. It ranges in [0, ∞) being 0 the perfect score. It has a singularity at $\overline{Obs} = 0$, which could occur for temperatures, for instance.

$$nMAE = \frac{MAE}{\overline{Obs}}$$

- NMAE: Mean Absolute Error (MAE), in units of the observed standard deviation. It ranges in [0, ∞) being 0 the perfect score. It has a singularity at $\sigma_{Obs} = 0$, which could occur for precipitation, for instance, in places where no rain is recorded.

$$NMAE = \frac{MAE}{\sigma_{Obs}}$$

- RMSE: Root Mean Square Error. It measures the average magnitude of the forecast errors, weighted according to the square of the error. This score ranges in [0, ∞) being 0 the perfect score.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (For_i - Obs_i)^2}$$

  It does not indicate the direction of the deviations. The RMSE puts greater influence on large errors than on smaller errors, which may be appropriate if large errors are especially undesirable. However, it may also encourage conservative forecasting.

- nRMSE: Root Mean Square Error, in units of the observed mean. This score ranges in [0, ∞) being 0 the perfect score. As the NMAE, the nRMSE has a singularity at $\overline{Obs} = 0$, which could occur for temperatures, for instance.

$$nRMSE = \frac{RMSE}{\overline{Obs}}$$

- NRMSE: Root Mean Square Error, in units of the observed standard deviation. This score ranges in [0, ∞) being 0 the perfect score. As the NMAE, the NRMSE has a singularity at $\sigma_{Obs} = 0$, which could occur for precipitation, for instance, in places where no rain is recorded.

$$NRMSE = \frac{RMSE}{\sigma_{Obs}}$$

3.3. Distributional Similarity

The study of the distributional similarity is other of the aspects that describe the quality of a forecast. It could be considered as the average agreement between forecasts and observations, i. e., their similarity in climatological terms. Distributional similarity should be carefully examined, especially in climate change studies. These are the scores included under this label:

- Ratio: Ratio of wet days. Ratio between forecasted and observed frequencies of wet days. It ranges in [0, ∞) being 1 the perfect score.

$$Ratio = \frac{P(For=1)}{P(Obs=1)}$$

Note that this score is specific for the case of daily precipitation. The threshold considered for defining wet days is 0.1 mm. It presents a singularity when $P(Obs = 1) = 0$ (all days being dry).

- Bias: Additive Bias. It measures the averaged forecast error, being 0 the perfect or desirable value.

$$\text{Bias} = \frac{1}{n}\sum_{i=1}^{n}(For_i - Obs_i)$$

This score does not measure the correspondence between forecasts and observations, i. e., it is possible to get a perfect score for a wrong forecast if errors are compensated.

- NBias: Bias in units of the observed standard deviation. It ranges in [0, ∞), being 0 the perfect score. It has a singularity at $\sigma_{Obs} = 0$, which could occur for precipitation, for instance, in places where no rain is recorded.

$$\text{NBias} = \frac{Bias}{\sigma_{Obs}}$$

- RV: Ratio of Variances. It measures the ratio between forecast and observed variances, in units of the observed one. It ranges in [0, ∞), being 1 the perfect score. It has a singularity at $\sigma_{Obs} = 0$, which could occur for precipitation, for instance, in places where no rain is recorded.

$$RV = \frac{\sigma^2_{For}}{\sigma^2_{Obs}}$$

- KS p-value: p-value from the two-sample Kolmogorov-Smirnov test. It ranges in [0, 1]. The decision to reject the null hypothesis of equality of distributions occurs when the significance level equals or exceeds this p-value.

The Kolmogorov-Smirnov test for two samples of sizes n and n' respectively, measures a distance, $D_{n,n'}$, between both cumulative density functions. $D_{n,n'}$ is calculated as:

$$D_{n,n'} = \sup_x |F_{1,n}(x) - F_{2,n'}(x)|$$

where $F_{1,n}$ and $F_{2,n'}$ are the empirical cumulative distribution functions of the first and second sample, respectively. This test is one of the most useful and general nonparametric methods for comparing two samples, as it is sensitive to differences in both location and shape parameters of the empirical cumulative distribution functions. Therefore, this score should be considered for validation, especially when projecting under climate change scenarios.

- KSX p-value: p-value from the two-sample Kolmogorov-Smirnov test restricted to observations and forecasts under their respective X^{th} percentiles. It ranges in [0, 1]. The decision to reject the null hypothesis of equality of distributions occurs when the significance level equals or exceeds this p-value.

- PDF Score: It measures the overlap between observed and forecasted empirical probability density functions. This score ranges in [0, 1] being 1 the perfect score. This score is calculated as:

$$PDFScore = \sum_{i=1}^{200}(PDF_{For_i} - PDF_{Obs_i})$$

being the PDF_{For_i} the forecast probability density function and the PDF_{Obs_i} the observed probability density function both for the i^{th} bin. 200 discrete bins (classes) are defined for the whole range of observations and predictions. Then, the probability density for each class is estimated by Kernel Density Smoothing. Observed and forecast probability densities are then compared for each class, retaining the minimum of each pair. The resulting sample of minimum is finally summed up.

In the portal, Gaussian Kernels and a width parameter optimization for normal distributions are considered to estimate the probability densities. Therefore, the user must be aware that this score is more appropriate for validating temperature than for precipitation. In addition, the PDF Score is hardly sensitive to failures in the tails of the distributions. Thus, the user should not rely exclusively on this score for validation, especially when projecting under climate change scenarios. We strongly recommend to consider both KS and PDF scores in conjunction.

Note that, for the special case of daily precipitation, and due to the high mass of probability density located at zero, the KS-pValue, KSX-pVvalue and the PDF Score are calculated for the continuous part of the distributions, by considering exclusively the observed and forecasted wet days. The discrete event occurrence/non occurrence is validated through the above explained HIR, FAR and Ratio scores. For the 10-daily precipitation and the temperature at both time-scales, the latter scores are calculated over the entire observed and forecasted time series.

4. Bibliography

- Bedia, J., S. Herrera, J. M. Gutiérrez, G. Zavala, I. R. Urbieta, and J. M. Moreno, 2012. *Sensitivity of fire weather index to different reanalysis products in the Iberian Peninsula*. Nat. Hazards Earth Syst. Sci., 12, 1-10.
- Bohren, C. F., and B. A. Albrecht, 1998. *Atmospheric Thermodynamics*. Oxford University Press, New York, NY.
- Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A.J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J. J. Morcrette, B.K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J. N. Thépaut, F. and Vitart, F., 2011. *The ERA-Interim reanalysis: configuration and performance of the data assimilation system*. Quarterly Journal of the Royal Meteorological Society. 137: 553-597.
- Gutiérrez, J. M., D. San-Martín, A.S. Cofiño, S. Herrera, and R. Manzanas, 2011. *User guide of the ENSEMBLES downscaling portal.* Technical Note 2/2011. Santander Meteorology Group. Santander.
- Jollife I. T. and D. B. Stephenson, 2003. *Forecast verification*. Wiley. 240 pp.
- Stocks, B.J., B.D. Lawson, M.E. Alexander, C.E. Van Wagner, R.S. McAlpine, T.J. Lynham and D.E. Dube, 1989. *The Canadian Forest Fire Danger Rating System: an Overview*. Forestry Chronicle 65: 450-457.
- Von Storch, H., E. Zorita and U. Cubasch, 1993. *Downscaling of climate change estimates to regional scales: An application to winter rainfall in the Iberian Peninsula*. Journal of Climate 6: 1161-1171.
- Von Storch, H. and F. W. Zwiers, 2001. *Statistical Analysis in Climate Research*. Cambridge University Press. 484 pp.
- Wallace, J. M. and P. V. Hobbs, 1977. *Atmospheric Science: An introductory survey.* Academic Press, New York.
- G. P. Weedon, S. Gomes, P. Viterbo, W. J. Shuttleworth, E. Blyth, H. Österle, J. C. Adam, N. Bellouin, O. Boucher, M. Best, 2011. *Creation of the WATCH Forcing Data and its use to assess global and regional reference crop evaporation over land during the twentieth century.* Journal of Hydrometeorology, 12: 823-848.
- Wigley, T.M.L., P.D. Jones, K.R. Briffa and G. Smith, 1990. *Obtaining sub-grid scale information from coarse-resolution General Circulation Model output.* Journal of Geophysical Research 92, 1943-1953