Project no. GOCE-CT-2003-505539

Project acronym: ENSEMBLES

Project title: ENSEMBLE-based Predictions of Climate Changes and their Impacts

Instrument: Integrated Project

Thematic Priority: Global Change and Ecosystems

D2B.19 Extension of the ENSEMBLES web-based service for downscaling (Task 2B.2.9)

Due date of deliverable: Month 30 (28-Feb-2007)
Actual submission date: 13-June-2007

Start date of project: 1 September 2004 Duration: 60 Months

Organisation name of lead contractor for this deliverable:
Partner 62 UNIVERSITY OF CANTABRIA (UC)

Revision [14-May-2007
13-June-2007]

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	PU
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the Consortium (including the Commission Services)	

1. Introduction

This deliverable describes the advances achieved by University de Cantabria (UC) in the development of the ENSEMBLES' statistical downscaling portal during the last year (project month 18-30 period). A first prototype of the portal was already described in the month 18 deliverable D2B.4 "A first prototype of the web service for downscaling at seasonal-to-decadal timescales" (Task 2B.2.2). The main goal of this first version was to explore the needs of a reduced number of users from different sectors (agriculture, energy, etc.) in order to include their feedback and needs in the final design of the portal. However, this initial prototype had several limitations, such as the restricted geographical area (the Iberian peninsula), and the limited number of computing resources (the portal ran in a single machine). Therefore, a significant amount of funded and unfunded human resources (17.5 PM) was used for months 18-30 in order to extend the downscaling portal taking into account the known limitations and the users' feedback. During this period, significant advances have been achieved, allowing the extension of the portal capabilities in the form described in the following sections (the new version of the portal can be accessed through the following web site: http://www.meteo.unican.es/ensembles).

2. Feedback from End-User Communities

From the very beginning, the downscaling portal was designed to be a useful and friendly service for end-users with limited experience in the technical issues involved in numerical modelling: binary data formats, protocols for data access, etc. Therefore, the feedback from end-users was a key factor to improve and extend the downscaling portal. Different potential end-users (about a dozen) were contacted and invited to "test" the prototype portal with a real case study. Finally, two users with different characteristics were selected (both are ENSEMBLES partners):

- Electricité de France (EDF), contact person *Laurent Dubus*:
  The aim of this user was to downscale DEMETER and ENSEMBLES S2D (seasonal to decadal) hindcasts to get *daily maximum and minimum temperatures* for a small network of stations (about a dozen) to make electricity demand forecasts.

- Joint Research Centre (JRC), contact person *Fabio Micale*:
  The aim was to downscale S2D hindcasts to get daily *precipitation, radiation, wind speed, and maximum/minimum temperatures* to perform crop yield modeling in Europe using the JRC dataset of gridded observations with a resolution of 0.5º (hundreds of locations).

In both cases, the final goal was to compare the downscaled data with the direct GCM outputs and to estimate seasonal predictability.

The above users tested the portal for several months and provided a lot of feedback about limitations and bugs they found using the portal. Some of the most important comments, as well as the improvements undertaken to address such issues, are described below:

- *The need to upload and update their own data into the portal in a simple form.* For instance, EDF needs to use its own non-public dataset (with reserved rights) to perform downscaling. This option was reported to be unfriendly in the first prototype and, finally we formatted and uploaded manually the necessary data so they could start the downscaling task.

  In the extended version of the portal, a user-friendly uploading panel has been included to facilitate this task (see Figure 1).

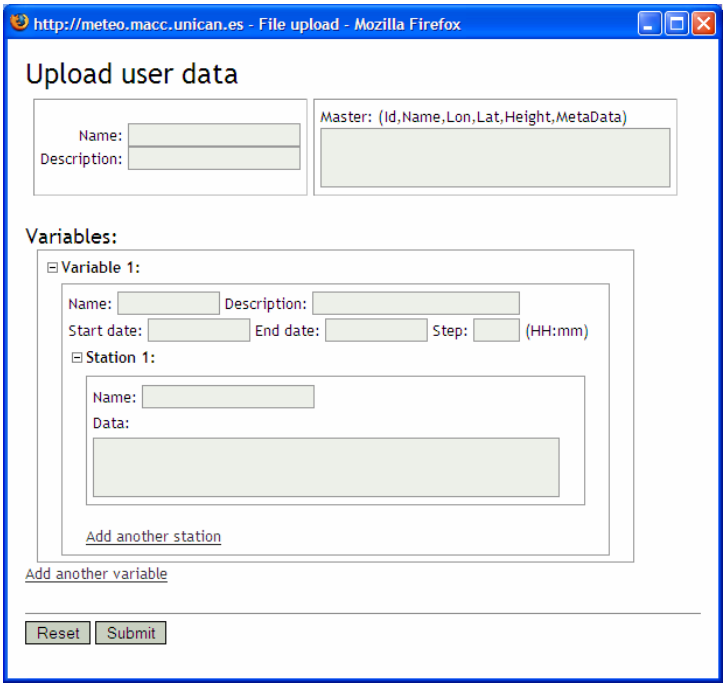

Figure 1. Panel for private observation data uploading and management.

- *The possibility to run perfect-model experiments (with reanalysis input predictors), obtaining an objective validation of the downscaling methods.* This would allow users to test different downscaling methods and to choose the most appropriate ones for their particular application.

  The extended version of the portal includes the possibility to run and validate the downscaling methods using reanalysis data. We want to remark here that this was an intensive computational task and a lot of effort has been dedicated to optimize the involved algorithms so it can run in a reasonable time (a few minutes). Figure 2 shows the new downscaling management window which allows performing cross-validation for the downscaling methods.

- *The convenience of running several downscaling tasks simultaneously*. In the first version of the portal, a downscaling job was defined as a single period of time for a single model (e.g., downscaling the output of the ECMWF's nine ensemble members for the December-February season using the November 1997 initialization).

In the extended version of the portal, the jobs are sent to the queue of a local cluster where they are scheduled and run independently of the portal. This migration of the portal computing services to a cluster gives flexibility and allows massive downscaling tasks to be submitted in a simple form. For instance, Figure 3 illustrates the possibility of selecting and submitting 48 different downscaling tasks as a single job (all the possible outputs of the ENSEMBLES S2D stream 1 simulation: 4 models x 12 initialization dates).

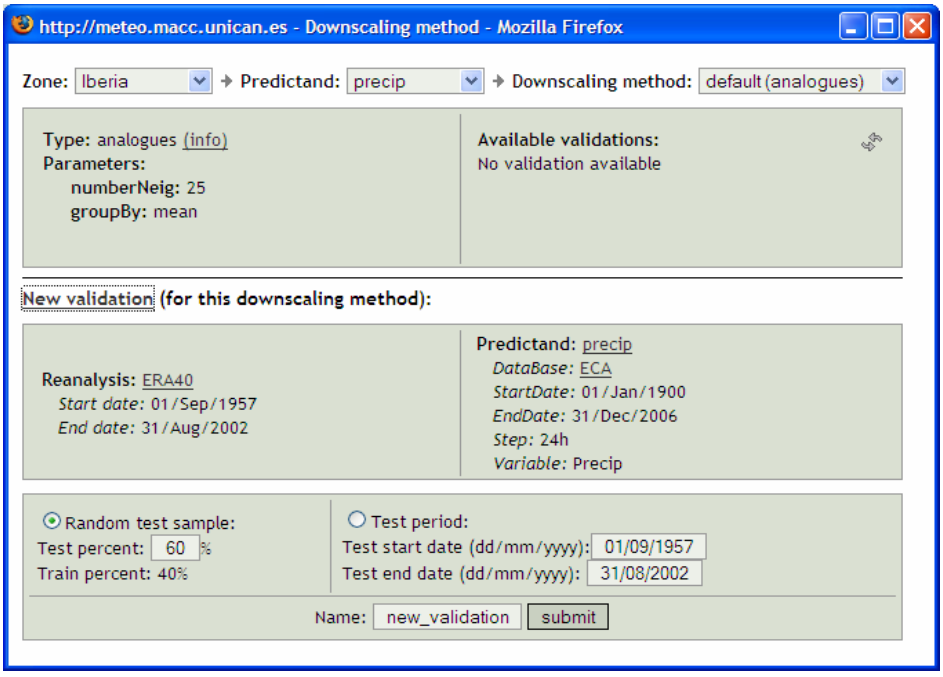

Figure 2. Downscaling methods management window. This allows the user to define and validate different downscaling methods.

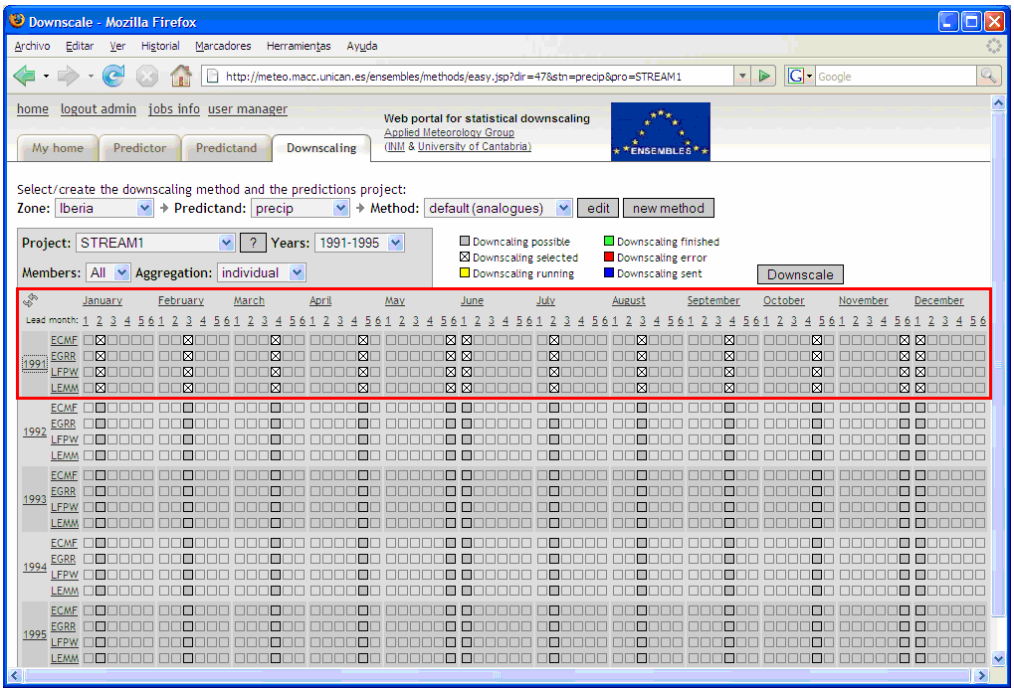

Figure 3. Downscaling window to select and submit downscaling jobs.

3. Extensions of the Downscaling Portal

In the new version of the portal (release 2), different extensions have been undertaken regarding the geographical area of study, the model outputs available to perform downscaling, the computing resources, the management of tasks, etc. In the following section, we briefly describe the most important extensions.

3.1. Extension of the Geographical Area

Initially, the Iberian peninsula was selected as the testbed area to run the case studies. However, it quickly turned out that this was a hard limitation to end-users so the area was extended to cover Europe. Both ERA40 and NCEP daily reanalysis fields are available over this area (see Figure 4). Moreover, observations for a number of local sites from the ECA network and for the JRC gridded observed data can be used to run downscaling experiments (the raw data itself is not directly available from the portal).

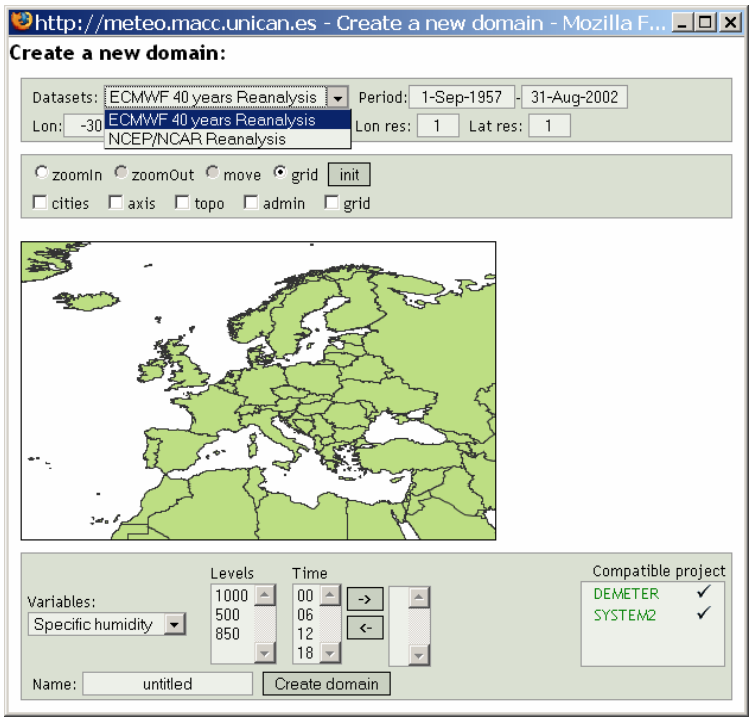

Figure 4. Geographical area available to run downscaling tasks.

Discussions with RT6 partners have started to explore the extension of the service to other world regions (i.e. South America, West Africa, India, etc.) depending on the availability of observations. In particular, a proposal to extend the service to the West African region was discussed during the AMMA/ENSEMBLES workshop held in Bamako 26 February to 2 March 2007 and was discussed further during the seasonal forecasting ENSEMBLES meeting in Barcelona 7/8 June 2007 (http://www.ecmwf.int/research/EU_projects/ENSEMBLES/amma/index.html).

3.2. Extension of the Statistical Downscaling Methods

In addition to the analog-like clustering algorithm included in the first version of the portal, a package of standard downscaling methods has now been included in the portal. This package is based on the MeteoLab Machine Learning Matlab Toolbox (see http://www.meteo.unican.es) and the available methods have been divided into three standard categories (weather typing, regression, and weather generator, as shown in Figure 5). These methods will be soon documented in the portal user guide.

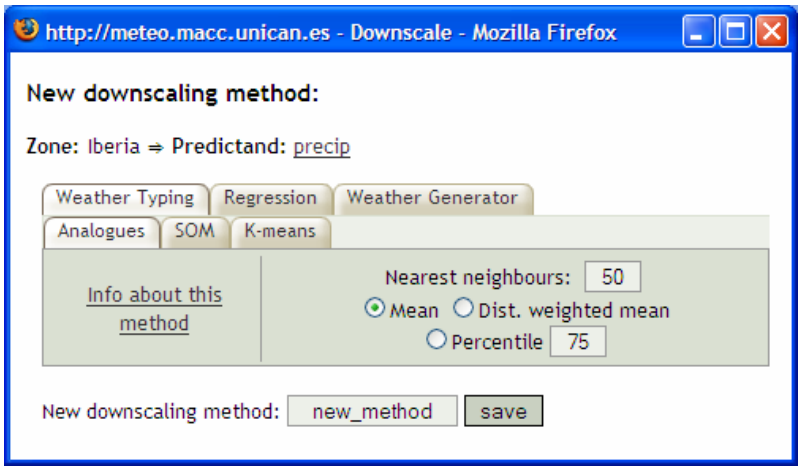

Figure 5. The "downcaling methods" panel allows the user to select different downscaling methods and appropriate parameters.

Discussion with RT2B statistical downscaling groups will take place to study the requirements for including in the web service additional statistical downscaling tools (i.e., in addition to the methods currently included in the portal).

3.3. Extension of the Computing Architecture

The first prototype was based on a simple computational architecture running on a single machine. This posed strong constraints to the type of tasks which could be performed by users (each simple downscaling job is an intensive computational task). Therefore, most of the effort during this last year has been devoted to avoiding this bottleneck by migrating to a decentralized computational architecture where multiple jobs from different users can run simultaneously. In the extended version, the computing is performed in a local cluster with GRID computing capabilities (therefore, the aggregation of new computing and storage resources will be an easy task). Figure 6 illustrates this new computational scheme; more details about this new architecture are given in Cofiño, San-Martín and Gutiérrez (2007).

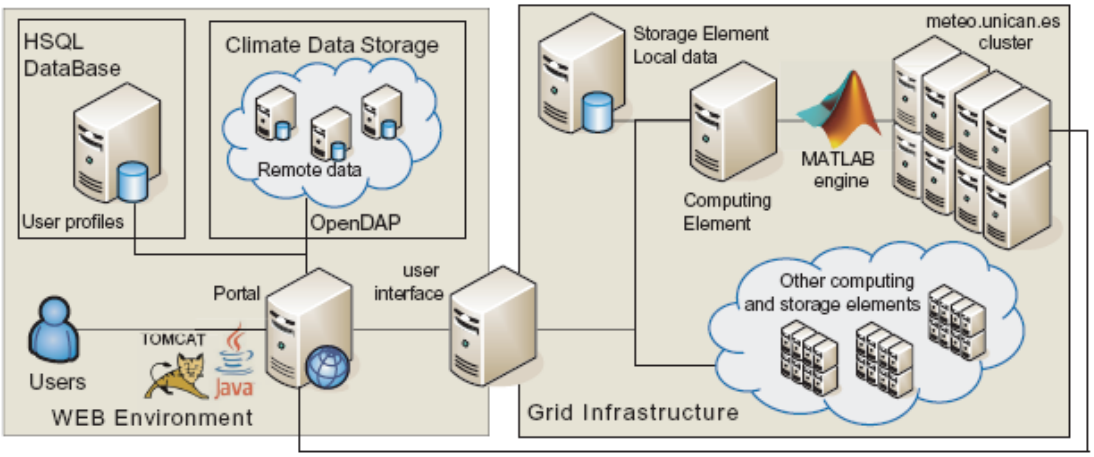

Figure 6. New computing architecture of the downscaling portal.

3.4. Extension of the Data Management

In the first prototype, both model and observed data were stored locally in the downscaling portal. Thus, only a reduced subset of information was available for users. The extended version of the portal includes an OpenDAP client to connect online to remote servers to obtain the model output data requested by the user in a typical session. DEMETER and ENSEMBLES seasonal-to-decadal model outputs are stored at ECMWF in OpenDAP servers, so these databases will be available online in the portal. Moreover, the climate change regional simulations will be also stored in an OpenDAP server at DMI. However, global simulations of climate change will be stored at the CERA server which does not include this possibility (see Section 4). We have analyzed the possibility to build a bridge to the java API used at CERA but at this time there is no possibility to include this data online in the portal.

Figure 7 shows the present (left) and future (right) data management structures of the portal. The migration to the OpenDAP client is work in progress and it is expected to be finished by the workshop on seasonal prediction (World Climate Research Program) to be held in Barcelona, Spain, June 4-7. A description of the portal will be presented in this workshop.

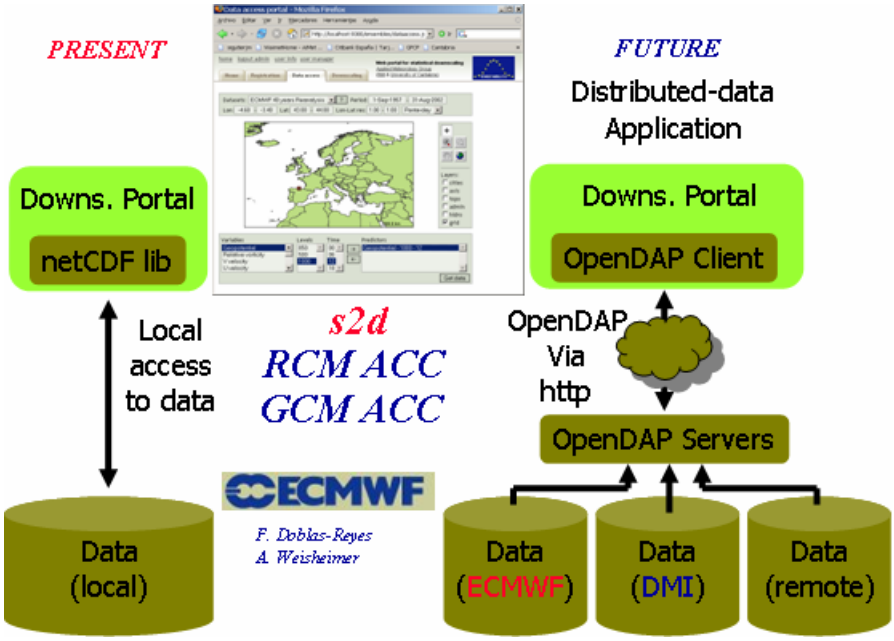

Figure 7. Data management structure of the downscaling portal.

3.5. Administration and Monitoring Capabilities

In the extended version of the portal, a new interface has been designed so the user can check the existing downscaling projects, edit its profile and monitor all the finished and pending tasks obtaining information about their status. To this aim, a *MyHome* initial tab has been added to the portal including all this information. This new feature provides a more flexible navigation.

For instance, Figure 8 shows the new initial window for a typical user, providing the information about the different available zones (with the corresponding downscaling tasks), the profile and submitted jobs.

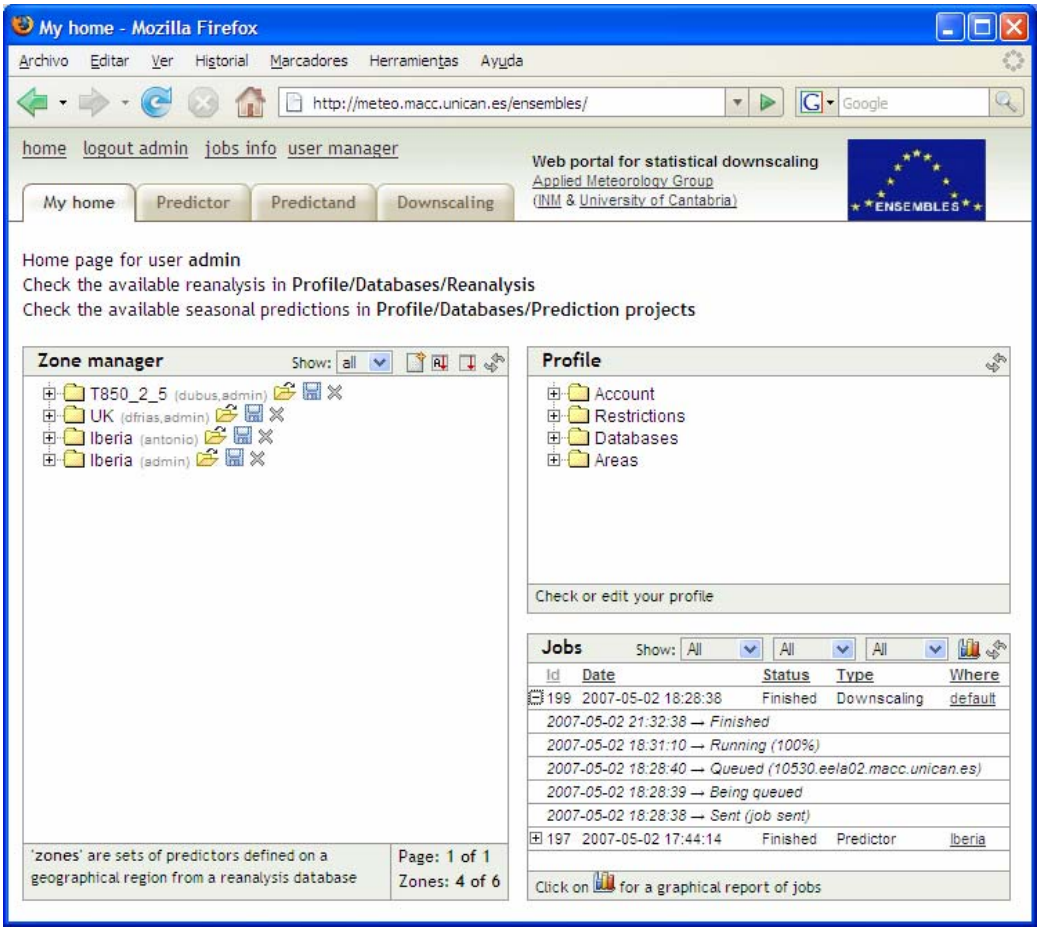

Figure 8. New interface with users (*MyHome* tab) providing the information about the zones, profile and jobs corresponding to the user.

4. Extension to Climate Change Timescales

During this one-year period, an analysis of the requirements for extending the service to climate change timescales and scenarios has started. The main bottlenecks found in the analysis were the storage of the global simulations in a system with no OpenDAP capabilities and the increasing computing resources needed to downscale climate model outputs for decades. The second problem has been already addressed with the new computing architecture of the portal (Section 3.3). The first problem is still pending and no solution has been adopted yet.

Finally, the recommendations to adapt the downscaling methods to climate change timescales are being evaluated in order to perform the necessary changes in the portal downscaling algorithms.

5. References:

A.S. Cofiño, D. San-Martín, and J.M. Gutiérrez (2007) A Web Portal for Regional Projection of Weather Forecast Using GRID Middleware. *Lecture Notes in Computer Science*, **4489**, 82–89.