Project of strategic interest NEXTDATA

Scientific Report for the reference period **01/01/2012-31/12/2012**

WP 2.5 - Archive of numerical simulations and projections (Resp: Silvio Gualdi, CMCC)

Partners: CASPUR, CMCC, CNR-ISAC, ENEA, ICTP

1. Scheduled activities, expected results and milestones

- Census of the global and regional numerical simulations and reanalyses available from the centres participating in the project. Harmonization of storage protocols of numerical data and of the procedure for data access and tranfer. Start of the archiving activities of the available numerical simulations.

- Development of collaborations and relationships with international projects and initiatives aimed at sharing and disseminating climate data from numerical simulations, such as CMIP5, CORDEX, ENSEMBLES, ECEarth, DRIHM and HyMex. Discussion with the scientific community and with the major European programs performing climate simulations, in order to define strategies for the use of the data.

- Definition of the "scientific quesitions" at the basis of the WP coordinated experiments aimed at producing high-quality data for the investigation of physical and dynamical climate processes of relevance for the areas of interest of the project (mountain areas with complex orography and the Mediterranean basin).

- Design and implementation of specific numerical experiments to be carried out during the project and attainement of the first results from new global and regional simulations for the areas of interest. Organization of two meetings of the researchers involved in the project, aimed at designing the climate numerical simulations, which will be made available in the project archives.

Milestone M2.5.1 (Month 12): Completion of the census of climate simulations and harmonization of storage protocols; definition of "scientific questions" and simulation strategies.

2. Deliverables expected for the reference period

- Deliverable D2.5.1: Report on a census of available climate simulations
- Deliverable D2.5.1: Report on the scientific questions

3. Activities conducted during the reference period

3.1 Research activities

During the first year of activity, all the partners of the WP have contributed to the census of the global and regional numerical simulations and reanalysis made available by the partners of the project, as required by the Project Executive Plan. In particular, information was collected about the available data, the storage requirements, the size and the description of the individual files, in order to harmonize the appropriate storage protocols and procedures to access the data.

Efforts have been started in order to establish collaborations with international programs and projects that involve the production of climate data by means of numerical simulations, such as, for example, CMIP5, CORDEX and Med-CORDEX, HyMex, ENSEMBLES, EC-Earth and DRIHM. These collaborations will be aimed at defining strategies and policies of use and dissemination of the available data.

Five WP meetings with all of the partners, have been held. During these meetings the main issues concerning the project activities have been discussed, such as, for example, the definition of the criteria for the census of the numerical simulations and reanalysis available from the partners, the harmonization of storage protocols of numerical data and of the procedure of data access and tranfer, the experimental set up of the coordinated integrations. A first meeting was held on the 26th of March 2012, during which the general guidelines for the definition of the characteristics of the experiments have been discussed, together with the scientific issues which the experiments will help to explore. Specifically, the following set of experiments have been identified as those of interest for the project:

- downscaling of climate simulations for the Alpine region (with a comparison of the dynamical with the statistical and stochastic techniques);
- simulations of the Mediterranean regional climate (past, present and future projections) performed with high resolution can coupled models able to reproduce the small-scale features and processes that characterize the basin;
- climate simulations for regions with complex orography (eg. Andes), aimed at identifying which spatial resolution is necessary to correctly reproduce the main features of the observed winds and of the atmospheric circulation in the area.

The second WP meeting was held in electronic form (teleconference) on May 31, 2012, in order to discuss and define the guidelines for the collection of information on the census of the climate simulations available at the centers partner of the project, subject of the deliverable D2.5.1 ("Report on the census of climate simulations").

In the third meeting (July 4, 2012), the construction of a network of data portals to be implemented within the centres partners of the WP has been discussed. The network of data portals will allow the sharing of the climate simulations and of metadata information, making them available and accessible to the scientific and user communities. This network of portals will be the backbone of the data disseminition system of the project. In particular, during the meeting, the partners discussed and decided the main technical features of the data archives (eg, type of servers, basic format of the data, storage protocols and mode of access), as a basis for the “general portal” of the project.

In the fourth WP meeting, which took place on September 14, 2012, the partners have reported on the status of project activities at each center, updating, in particular, about the advances in the activity of the census of simulations (Deliverable D2.5.1) and issues related to the implementation of servers with similar characteristics and shared access protocols to the data.

During the fifth meeting, held on 6 November 2012, the partners discussed and defined the details of the "scientific questions" and the related coordinated numerical experiments to be conducted within the WP. During the meeting, the partners also reported on the results of preparatory studies made in order to explore issues and possible criticalities related to the coordinated experiments. The results of these discussions have contributed to define the content of the deliverable D.2.5.2 ("Report on the" scientific questions ").

Finally, during the first year, all partners have completed a series of experiments and climate simulations aimed at the completion of the data sets to be made available to the scientific and users communities, and to the definition of the main scientific issues to be investigated in the project. For example, CNR-ISAC and CMCC completed, with their climate models, the simulations (long-term projections and decadal predictions) planned withing the CMIP5 program activities. ICTP and ENEA, on the other hand, have completed with their regional models the downscaling simulations planned within the CORDEX and Med-CORDEX programmes. While CMCC and ENEA, have conducted preliminary studies to identify the problems and difficulties related to the production of long climate simulations for the Mediterranean region, aimed at reproducing the variability observed in this area in the past decades.

3.2 Applications, technological and computational developments

At **ICTP**, substantial development was carried out for the RegCM regional model, including both enhancements in the physics schemes (convection, land surface, radiation, planetary boundary layer) and code optimization. In particular a version of the model was developed which considerably improves parallelization and scaling compared to the previous one. This version will allow us to increase the model resolution while keeping reasonable computation times.

CMCC has developed a regional coupled ocean-atmosphere model of the Mediterranean region, composed by the limited area atmospheric COSMO model and by the NEMO-MFS model of the Mediterranean Sea. The atmospheric component is implemented with a 25 km resolution, whereas the ocean component has a horizontal resolution of about 6.7 km and 71 vertical levels. Furthermore, preliminary tests have been conducted to assess the ability of the model to reproduce the observed climate features of this area. The coupled model will be used to make the climate simulations for the Mediterranean area (including the Alpine region) planned within the framework of the project.

The raw outputs of the climate simulations performed in the frame work of the CMIP5 programme with EC-Earth **(CNR-ISAC)** and with the **CMCC** climate models (CMCC-CM, CMCC-CSM e CMCC-CESM) have been post-processed and a subset of variables (following CMIP5 recommended output fields and time frequencies) have been converted to CMOR2 netcdf format, enriching the netcdf files with ample metadata information. Additionally, a set of variables of interest for application to the study of snow cover and of precipitation and temperature extremes, not included in the CMIP5 archives, have been extracted at high temporal resolution (3hr).

The output of existing climate simulations performed by CNR-ISAC at CASPUR using the ECHAM-HAM 5.5 model, including the aerosol dynamics and transport module HAM2, have been collected and made available in a centralized archive at CNR-ISAC. Furthermore, also in collaboration with CASPUR, CNR-ISAC has implemented the new EC-Earth release v3 (released in October) on the CINECA/Caspur Matrix cluster. This version of the model is based on IFS 36r4, the latest version of the NEMO model (3.3.1) and the LIM3 sea-ice model. It will be run at T255 resolution (about 80 km) with 91 vertical levels. A series of successful test runs and scaling tests has already been performed on the Matrix cluster (http://www.to.isac.cnr.it/ecearth/ecearth3/scaling.html).

In the framework of the WP downscaling activities, starting from the EC-Earth model outputs, specific boundary condition files have been prepared for the periods 1960-2005 (historical) and 2006-2050 (RCP 2.6, RCP 4.5, RCP 8.5) and transmitted to ICTP in order to be used for the production of model simulations with a hydrostatic regional climate model for the HKKH focus area. At CMCC the boundary conditions for a series of downscaling experiments to be conducted with the non-hydrostatic, high-resolution (~8 km), atmospheric model COSMO-CLM have been prepared for the Alpine area. Similarly, CNR-ISAC started the implementation and preparation of a series of extremely high-resolution and numerically intensive dynamical climate downscaling experiments using the non-hydrostatic model WRF. This activity is performed using computing resources provided by the Gauss programme at the German Supercomputing Center LRZ. The focus will be on regions where significant impacts of climate change are expected and where precipitation modelling is particularly challenging due to the formation of intense and complex meteorological structures or to rich and complex orography (The Hindu-Kush-Karakorum-Himalaya, the Alps and the Mediterranean area, and the Caribbean/South American area). To this purpose the latest version of the WRF model (3.4.1) has been integrated with a modification (CL-WRF), which allows flexible use of GHG scenarios and implemented on the SuperMUC machine at LRZ. A first test experiment has been performed at resolution of 0.22 degrees, using ERA-Interim boundary conditions, for a period of one month.

A modified version of a consolidated stochastic method, the RainFARM procedure developed at CNR-ISAC, has been applied to climate simulations produced by the PROTHEUS regional climate model, developed at the ENEA centre in the framework of a collaboration between ENEA and ICTP. The aim of this study is to investigate whether the precipitation produced by this RCM, stochastically downscaled with RainFARM, is able to reproduce the main statistical properties of the precipitation measured by a network of rain gauges located in north-western Italy. We considered 122 rain gauges located in the Piedmont and Valle d'Aosta regions, spanning the time period from 1958 to 2001 and located at different altitudes, ranging from about 100 m to about 2500 m above sea level. We applied RainFARM also to the large scale driver of the PROTHEUS RCM, the ERA40 global reanalyses to compare the skill of the stochastic downscaling procedure when applied to precipitation fields having different spatial resolution (PROTHEUS ~ 30km, ERA40 ~ 100km) and to emphasize the role of dynamical downscaling.

All of the WP partners (**CASPUR, CMCC, CNR-ISAC, ENEA, ICTP**) have installed THREDDS servers, in order to provide access of available data to the project participants and to prepare for the future indexing of data through a general portal. The servers implement the OpenDAP protocol, providing access to complete metadata information on the available gridded fields. According to the planned server design, the access to the data will be provided by means of a http fileserver and a subsetting tool will allow to extract chosen spatial and temporal ranges for selected variables. The technical details of the server implementation have been discussed

and decided by the WP2.5 partners during the project meetings.

As an example, in the following the data archiving and delivering system developed by CASPUR is illustrated, similar to the systems implemented also by other project partners. The CASPUR infrastructure includes a typical distributed-memory supercomputer based on traditional X86 processors and a system server for the intensive I/O services (and for the external access service, virtualization and DBMS) which are interfaced to a Enterprise-class storage subsystem. Figure 1 shows the block diagram of the entire hardware infrastructure.

This infrastructure of high-performance computing and storage allows to support all of the simulations provided in CASPUR and to start storing all of the datasets provided from simulations carried out elsewhere. The transfer of large amounts of data available elsewhere has greatly benefited from the research network GARR, in which CASPUR is one of the strategic access points. This synergy between CASPUR and GARR has provided and will guarantee users with a NextData available bandwidth, continuous and specialist support for the resolution of issues related to breaks, data transfer bottlenecks and massive datasets (of the order of terabytes).

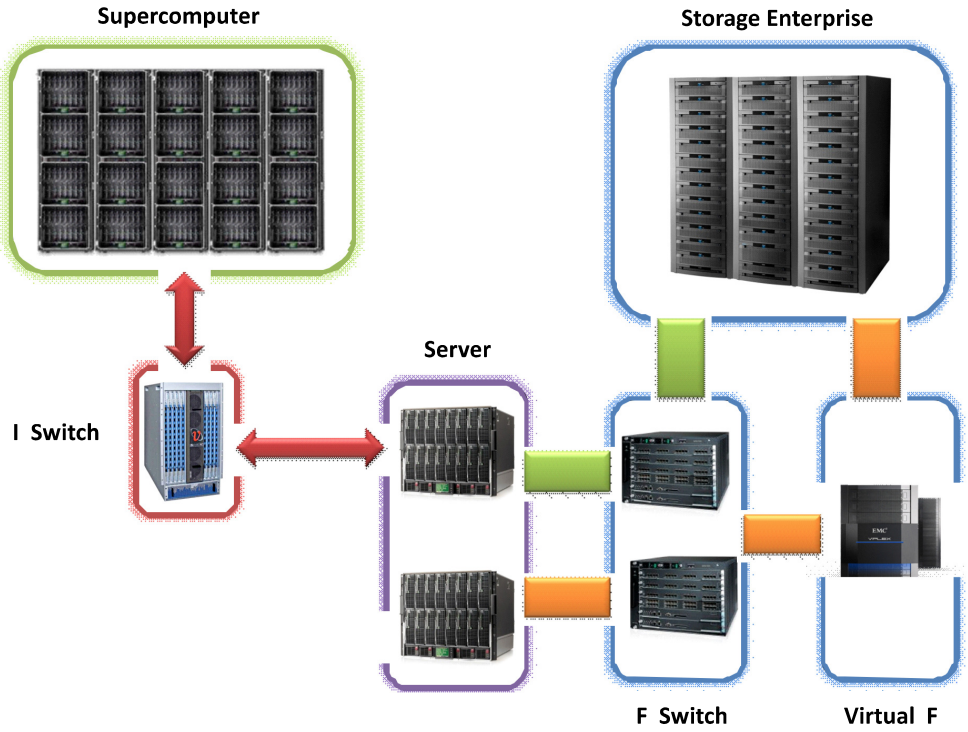

Figure 1. Block Diagram of the ICT / HPC for NextData.

To make the computational data available to the scientific community, it was decided to use the THREDDS Data Server (TDS) version 4.2 (then upgraded to 4.3), which is an acronym for Thematic Real-time Environmental Distributed Data Services tool. The TDS is a web server that provides access to data and metadata, using different, including OPeNDAP, OGC WMS and WCS, HTTP.

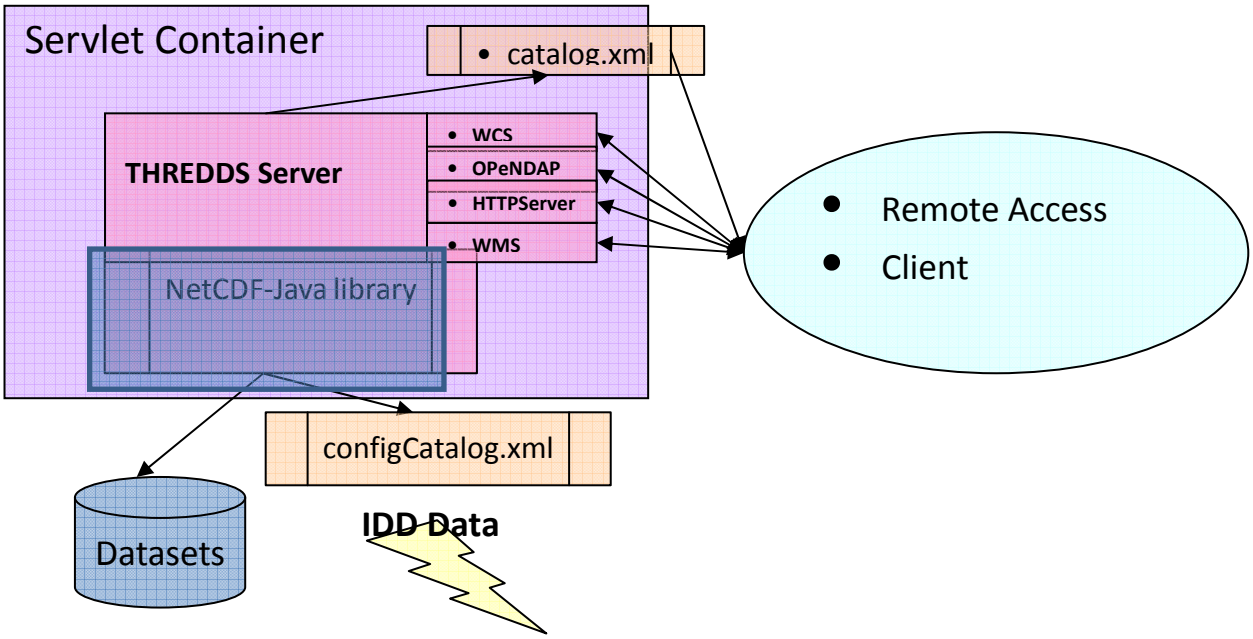

Figure 2- scheme of the implementation of a THREDDS Data Server (TDS)

The Thredds data server relies on the open source Apache Tomcat web server, which is an implementation of the Java Servlet and JavaServer Pages, and it provides a range of common and standard data access protocols (OPeNDAP, OGC WMS e WCS, HTTP). Figure 2 shows a typical implementation of a TDS system.

3.3 Formation

At CNR-ISAC, a Master Thesis (Laurea Magistrale) has been supervised, on the topics of validation of global EC-Earth precipitation against available satellite and gridded station archives, on the analysis of statistics of extreme precipitation in the historical period and on the role of mid-latitude disturbances on winter precipitation in the Karakorum and the relationship with teleconnection patterns.

Furthermore, at CNR-ISAC a first cycle degree thesis has been supervised on the comparison of precipitation and temperature produced with EC-Earth for the historical period with available station observations and on the analysis of trends and statistics of precipitation and temperature scenario simulations, for the area of North-Western Italy. During the same period, a post-graduate fellowship has been started on the development and use of stochastic downscaling techniques applied to climate models outputs.

At ICTP, a workshop was conducted in May 2012 on the latest version of the RegCM system, RegCM4. This workshop included both theoretical classes and tutorial/laboratory sessions. About 90 students participated to this workshop. In addition, 4 students, from Brazil, India, Mexico and Hungary, were invited to ICTP for a period of 6 months to carry out the simulations in 3.1 and start their analysis under the supervision of the scientific staff of the Centre.

3.4 Dissemination

During the first year of the project, several meetings were held on the use of stochastic downscaling techniques, with representatives of potential local users, such as, for example, ARPA Piemonte.

3.5 Conferences, Workshops, Meetings

Scientific results obtained with model simulations developed in the framework of this WP have been presented by CNR-ISAC in talks at the following workshops and conferences:

- International meeting with EC-Earth contributors/users in Reading, UK. May 30 - 31, 2012.
- ECRA (European Climate Research Alliance) workshop: "Changes in the hydrological cycle", CNR, Rome, 5-6 March 2012.
- Workshop: "Orographic Precipitation and Climate Change", NCAR, Boulder, 13-15 March 2012.
- European Geosciences Union (EGU) General Assembly 2012, Vienna, 22-27 April 2012.
- Workshop: "Contribution of science and cooperation to the sustainable development of the Central Karakorum National Park", Islamabad, 4-7 June 2012.
- 6th HYMEX Workshop, Primosten, Croazia, 7-10 May 2012.
- ECSAC 2012: CLIMATE CHANGE: marine and mountain ecosystems in the Mediterranean region, XII International Conference on Science, Arts and Culture, Veli Lošinj, Croazia, 27-30 August 2012.
- Conferenza MED-Clivar 2012: “The climate of the Mediterranean region: understanding its evolution and effects on environment and societies”, 26-28 September 2012

4. Results obtained during the reference period

4.1 Specific results (Data libraries, Measurements, Numerical simulations, etc)

A large amount of climate data (both simulations and re-analyses) has been produced, collected and organized during the first year of the project. The results of this work have been summarized and illustrated in the project deliverable D2.5.1 (Report on a census of available climate simulations). In addition the storage and transfer of available numerical data on the data servers implemented by the participants in the WP has started. These data servers will allow direct access to the simulation data and will provide the backbone infrastructure for the general portal of the project.

The downscaling activities have led to important, though preliminary, results. For example, the high-resolution precipitation fields obtained by downscaling the PROTHEUS output with RainFARM reproduce well the seasonality and the amplitude distributions of observed precipitation during most of the year, including the extreme events. RainFARM produces better results, compared to the observation, when it is applied to PROTHEUS than ERA40, highlighting the added value of dynamical downscaling. However, the biases of the regional climate model or of ERA40 cannot be corrected by this stochastic downscaling procedure: RainFARM can introduce variability at the scale which are not resolved by the physical models, but it cannot replace the physically-based models, also used to better understand rainfall dynamics.

4.2 Publications

- Parodi A., von Hardenberg J., Provenzale A., Emergence of large-scale patterns in moist atmospheric convection, submitted to Journal of Geophysical Reasearch – Atmospheres.
- D. D'Onofrio, E. Palazzi, J. von Hardenberg, A. Provenzale, V. Artale and S. Calmanti, 2012: Stochastic rainfall downscaling of a regional climate model over north-western Italy. To be submitted to Journal of Hydrometeorology. In preparation.

4.3 Availability of data and model outputs (format, type of library, etc)

A detailed list and description of the data made available to the project partners and users is provided in the project deliverable D2.5.1 (Report on a census of available climate simulations).

4.4 Completed deliverables

In the course of the first year project the partners have completed the project deliverables :
- Deliverable D2.5.1: Report on a census of available climate simulations
- Deliverable D2.5.1: Report on the scientific questions

as scheduled in the project implementation plan.

5. Differences between planned and performed activities/results/deliverables

No significant differences have occurred during the first year between the planned activities and those actually performed.

6. Expected activities for the following reference period

In the course of the second year of the project, all of the WP2.5 partners will contribute to the following activities:

- continue the production of global and regional numerical simulations targeted to regions of interest of the project, including specific experiments with dynamics of aerosol;

- complete the provision of the contents of the archive of numerical data, with a particular focus on the Mediterranean area, the Alpine region and the region HKKH;

- implement high-resolution numerical models at the local scale, not hydrostatic, for the simulation of the dynamics of climate and environment in mountain areas with complex orography;

- continue the work on stochastic downscaling to investigate future scenarios of precipitation at high resolution in north-western Italy.

- conduct a comparison between different techniques of dynamical, statistical and stochastic, downscaling in order to understand the features and applicability limits for specific case studies.

- generate high-resolution data archives, using downscaling techniques applied to output of observations and models at low resolution (both spatial and temporal).

- organize (at least two) meetings of the project researchers, in order to discuss climate scenario experiments, and the distribution and use of the numerical data.