



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

Deliverable D2A.4.3

Preliminary version of the ECMWF public data server for dissemination of seasonal-to-decadal simulations

Due date of deliverable: month 30

Actual submission date: month 30

Start date of project: 1 September 2004

Duration: 60 Months

Organisation name of lead contractor for this deliverable: ECMWF

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)	

ENSEMBLES

Deliverable D2A.4.3

Authors: Francisco J. Doblas-Reyes, Carlos Valiente, Manuel Fuentes
(ECMWF)

Due: month 30

Preliminary version of the ECMWF public data server for dissemination of seasonal-to-decadal simulations

1. Introduction

This report describes the release of the preliminary public data server for dissemination of seasonal-to-decadal ENSEMBLES simulations hosted at ECMWF. The service serves global hindcasts using a MARS-based web portal and a tool based on the OPeNDAP technology. At present, the daily and monthly mean data generated in the Stream 1 set of seasonal and annual simulations is available. Fields from the DEMETER multi-model dataset are also available. The system has been conceived and developed to help ENSEMBLES users and external scientists to access the hindcasts. The service is fully documented on-line.

2. The data server

The s2d ENSEMBLES global hindcasts are disseminated using a MARS-based tool and a server using the OPeNDAP technology. While the first tool offers a quick and easy way to interactively download the data, the second one offers the possibility of accessing the dataset to external client applications, which is expected to widen the use of the data. For instance, an application user with, let's say, a disease model, might be interested in downloading a single set of hindcasts to employ them as an extrinsic variable. This user might save time and effort employing the MARS-based system. As a very different example, an external application, such as the Climate Explorer, requires the access to a command-based system to retrieve the data automatically, which is a requirement easily satisfied by the OPeNDAP server.

The servers can be visited from the RT1 web site:

www.ecmwf.int/research/EU_projects/ENSEMBLES/data/data_dissemination.html

The OPeNDAP server is an aggregation server (also known as Thredds) that collects all the data (either daily data or monthly means) into a single dataset. Although many OPeNDAP servers act as a tool to aggregate remote datasets, this one accesses locally NetCDF data files archived in the ECMWF ENSEMBLES machine. OPeNDAP offers access to the data in an interoperable manner from client applications. Familiar data analysis and visualization applications can be used as clients: IDL, Matlab, Ferret, NetCDF operators, etc. OPeNDAP also allows a client to request any subset of the dataset. Visualization and data handling of the data can be carried out by either using the NetCDF/Java tool provided with OPeNDAP or by installing either ncBrowse or ODC in your computer.

The dissemination of s2d hindcasts using an OPeNDAP server requires a unified framework for product standardisation to provide a coherent service. The standardisation has been achieved by providing rules to encode multi-forecast system ensemble hindcasts in CF-compliant NetCDF files. This is the first time such an attempt has been undertaken. A document available from the URL above submitted for discussion to the CF group describes the structure of the metadata to be used with special emphasis on the dimensions of the file and the variables that describe the way the ensemble has been generated. Operational requirements have been taken into account at the time of designing the structure. A second document lists the standard names of the variables as required by the CF convention. These rules are expected to be approved by the CF group, following the discussions carried out in recent months. It is expected that multi-forecast system simulations carried out at monthly, seasonal, interannual and centennial time scales will be encoded in the future following similar guidelines.

The MARS-based server uses the technology developed in the DEMETER project. The server is user friendly, interactive and provides the data in GRIB and NetCDF (not CF-compliant though) formats. Specific areas can be easily selected and a tool to plot the fields is also available. Note that the system is intended for interactive access and the number of fields to be retrieved in one request is limited. Users with high or frequent demands of data are strongly encouraged to access the OPeNDAP server with a script.

3. Future work

We are considering several tasks to make the service more efficient and widen its scope:

- Archive and disseminate the stream 2 data as soon as the integrations have been made and the data quality controlled.
- Liaise with KNMI and Universidad de Cantabria to link the Thredds server with their Climate Explorer and Downscaling Web Portal services. This task is extremely important as it would increase the usability of the data by the application partners in RT6 and those external to the project.
- Investigate the capability of the Thredds server to aggregate NetCDF compressed files. Thredds does not seem able at present to properly aggregate 16-bit NetCDF files. Sorting out this inconvenient would cut in half the size of the dataset stored in the ENSEMBLES machine.

- While the ENSEMBLES machine has 5Tb of disk space available, at present two copies of the data are stored: one in GRIB format and another one in NetCDF. This is required by the differences of the MARS and OPeNDAP technologies. It might be possible in the future to create a convertor of GRIB to CF-compliant NetCDF compatible with OPeNDAP that could allow for conversions of data “on-the-fly”. The MARS-based server will still be available for its simplicity and convenience of use.

4. Acknowledgments

This work would not have been possible without the collaboration of many colleagues of the Operations Department at ECMWF.