Seventh Framework Programme
Theme 6 [SPACE]

Project: 607193 UERRA

Full project title:
Uncertainties in **E**nsembles of **R**egional **R**e-**A**nalyses

Deliverable D4.3
Data services and visualisation prototype

WP no:	4
WP leader:	Albert Klein Tank
Lead beneficiary for deliverable :	ECMWF
Name of author/contributors:	Richard Mladek, Manuel Fuentes
Nature:	Report
Dissemination level:	PU
Deliverable month:	12
Submission date: March 24, 2015	Version nr: 5

Report for Deliverable 4.3 (D4.3): Data services and visualisation prototype

By Richard Mladek and Manuel Fuentes

European Centre for Medium-range Weather Forecasts

For this deliverable, the goal is to work on future UERRA (Uncertainties in Ensembles of Regional Re-analysis) tools for data access and visualisation. Rather than developing yet another portal and other tools for access to and visualization of the ensemble reanalysis raw datasets from the scratch, the synergies with existing systems and systems under development have been exploited to the full. As the UERRA datasets are not fully to disposal yet the former similar EURO4M (The European Reanalysis and Observations For Monitoring) datasets are used at this stage as a data source for testing and preparation of all related tools.

Section 1: Building data portal prototype and related tools

1.1 Introduction

A well designed data portal and provision of user interfaces are crucial for the work in this project for creating products and interaction with users. ECMWF (European Centre for Medium-range Weather Forecasts) has a lot of experience in these areas, having been a partner in projects like TIGGE (The Observing System Research and Predictability Experiment), where ensemble forecast data from nine operational centres has been exchanged on a daily basis, archived and served using ECMWF systems since 2006. ECMWF led discussions about encoding data using a common format in order to enable homogeneous access for users. TIGGE data is currently available to users via a web data portal [1]. ECMWF was partner in another project GEOWOW (GEOSS Weather, Ocean and Water [1]) and is also a partner in currently ongoing project S2S (Sub-seasonal to Seasonal project [3]). In both of them the Centre has been archiving TIGGE-like data from different models run on various domains. Again, the common encoding has had to be found so data from different originating centres and models could be accessed and used in a seamless way. All this experience will help in shaping of data formats and encodings to be used for the data produced in UERRA. Many of the already existing tools for data access and visualization can also serve for UERRA purposes after appropriate adjustments and amendments.

1.2 UERRA website

A project website (related to data archiving in this context) has been established under a wiki type system run at ECMWF (Figure 1). This website will play a crucial role within the whole project to keep all information in one place well organized and available for partners.

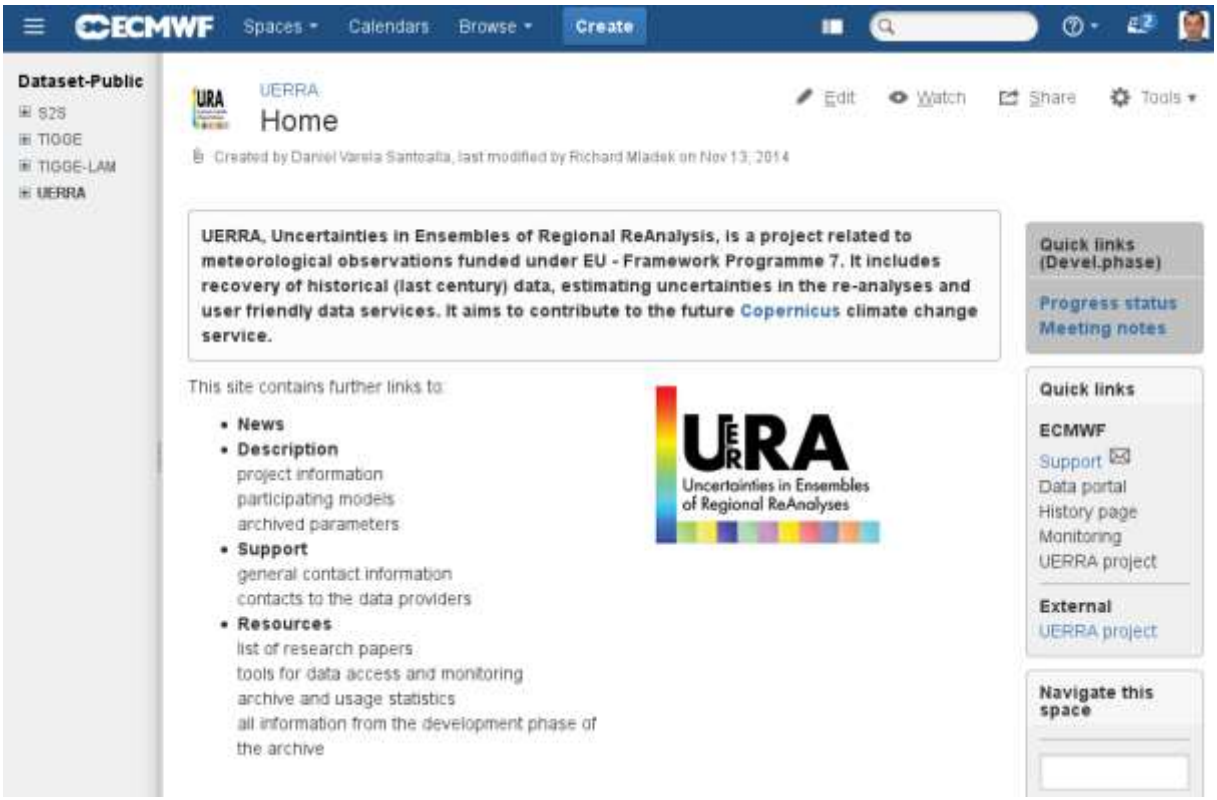

Figure 1. The UERRA website at ECMWF

1.3 EURO4M testbed

As the UERRA project has started only recently the work on reanalyses datasets by partners is still in progress. Because of this it was decided to use former EURO4M outputs as a testbed for preparation of for example data portals and all related data handling tools. Some EURO4M data (UM/4DVAR UK Met Office outputs) has been already available in MARS (Meteorological Archive and Retrieval System) at ECMWF and could be used directly for first testing. As the next step it was decided to archive more comprehensive available EURO4M data subsets for period 2008-9 from more partners. It was agreed to archive only selected single level parameters:

- 10 meter u-velocity
- 10 meter v-velocity
- Convective precipitation
- Large-scale precipitation
- Orography
- Snow depth water equivalent
- Snow fall water equivalent
- Surface air maximum temperature
- Surface air minimum temperature
- Surface air relative humidity
- Surface air temperature
- Time-integrated surface net solar radiation

- Time-integrated surface solar radiation downwards
- Total cloud cover
- Total precipitation

The data will be archived in its original encoding format (GRIB1) as it is without any further optimization (e.g. to achieve unified parameters' names and definitions). This simplified approach was chosen to speed up the data availability via MARS for verification purposes of UERRA WP3.

The list of expected EURO4M datasets with current progress status is shown in Table 1. As already mentioned the outputs from UK Met Office (MO) model called UM/4DVAR were archived in MARS in 2012 at the initiate of the UK Met Office (MO). The data from SMHI (Swedish Meteorological and Hydrological Institute) has been already provided for the whole period 2008-9 and its archiving in MARS has successfully started. The data samples from two remaining partners DWD (Deutscher Wetterdienst) and MF (Météo-France) were received and the first archiving tests have been done. After some necessary minor fixes of the data format the complete datasets for defined period will be asked and archived.

Model	**Data status**	**Next milestone**
COSMO (DWD)	Test data received	Archive in MARS
HIRLAM (SMHI)	Test data received	Archive in MARS
MESCAN (MF)	Test data received	Archive in MARS
UM/4DVAR (MO)	Archived in MARS	

Table 1. EURO4M datasets agreed as the testbed for UERRA

The data portal shown in Figure 2 was developed in the frame of UERRA for EURO4M dataset provided by UK Met Office. After a simple registration and agreement with data usage policy a user can retrieve required data by selecting desired combination of time period, parameters, forecast steps etc. This kind of data portal will not be provided for the other EURO4M datasets because of their expected inconsistencies (data is stored in its original format as stated above in this chapter).

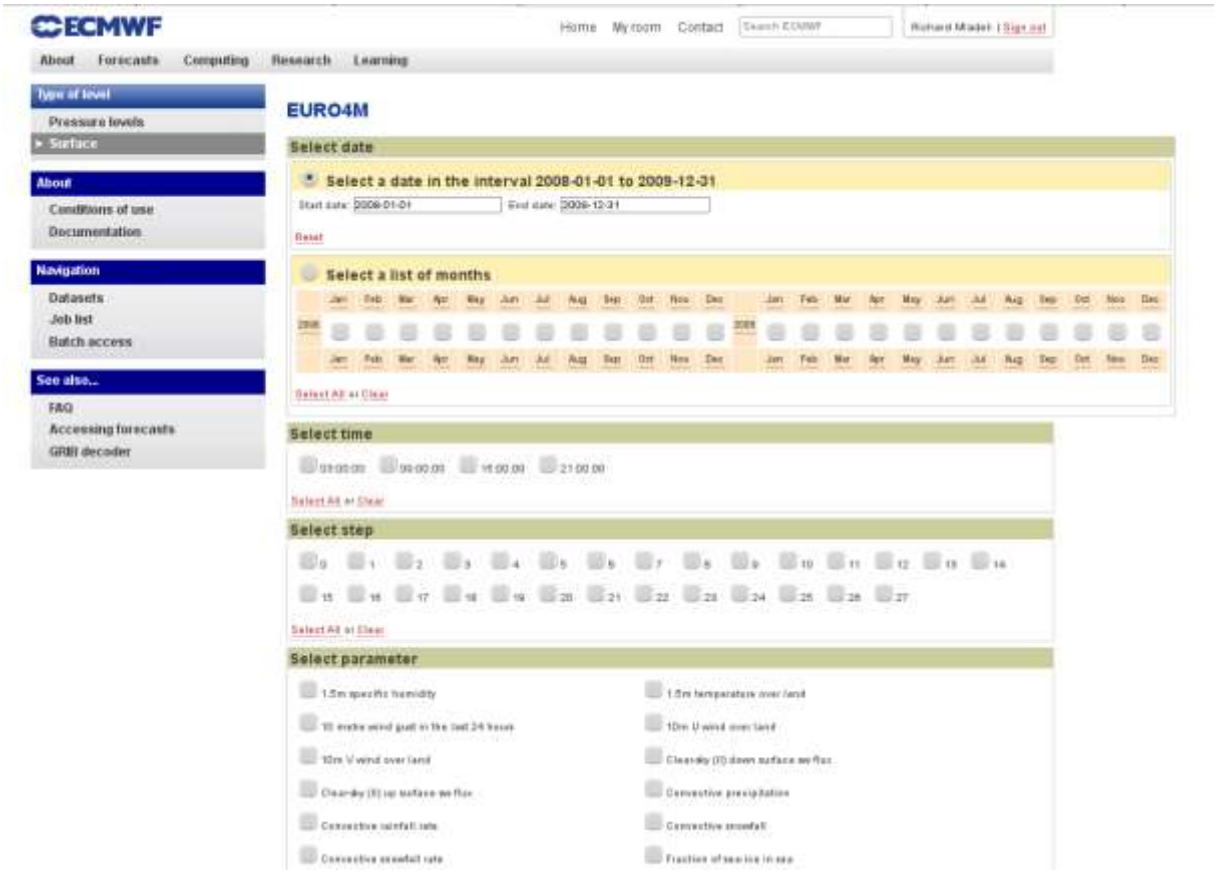

Figure 2. EURO4M data portal allowing easy access to UM/4DVAR model outputs

1.4 UERRA datasets

As the UERRA datasets are not ready yet, the most important step at this stage is to understand in detail what is going to be produced and required to be archived. From MARS archive point of view it is crucial to have gathered all possible information about future datasets and their parameters. Without exact description of each model set up and produced parameter it is not possible to create the best possible MARS set up for efficient access to the datasets. Any information not known in the planning period of MARS for UERRA set up could easily end with suboptimal archiving in the future or even impossibility to archive what is required at all. Two working documents describing datasets and parameters have been prepared by WP3. They are shared with others via google drive [4] and the work on their finalization is in progress. Many discussions and information exchanges are still expected to happen before reaching their final form.

In Table 2 there is a list of eight up-to-now identified UERRA datasets. As can be seen, the first sample data have been already obtained from two models so that the work on the next milestone, conversion to GRIB2, could start. It is one of the fundamental steps as the GRIB2 is the main archiving format for MARS and one of the most important meteorological data formats under WMO's (World Meteorological Organization) governance. Very important is to achieve homogenous GRIB2 encoding compliant with standardised WMO definitions which have been used for TIGGE, GEOWOW and S2S datasets as well. Thanks to that

approach the data can be connected in the future to all other available services using common standards for exchange of meteorological data and products.

Model	**Data status**	**Next milestone**
COSMO (MIUB)	Test data received	Convert to GRIB2
COSMO/En (MIUB)	Test data expected	
HARMONIE/V1 (SMHI)	Test data received	Convert to GRIB2
HARMONIE/V2 (SMHI)	Test data expected	
MESAN (SMHI)	Test data expected	
MESCAN (MF)	Test data expected	
UM/4DVAR (MO)	Test data expected	
UM/En4DVAR (MO)	Test data expected	

Table 2. List of identified datasets which should be produced within UERRA project and archived in MARS

In the same location on google drive [4] there is another at the moment even more important "living" document describing parameters required to be archived in MARS. The modelling systems producing UERRA reanalyses are very complex ones combining the latest knowledge from many areas of numerical weather prediction (ensemble forecasting, data assimilation, physical parameterizations etc.). The constituent models/analyses are also heterogeneous. For example most data is from 3-D analyses and forecast models with some deterministic or control runs but most have ensemble members too. Further the models are different scientifically and technically, hence not all their variables and output parameters comply with each other. Then there are the surface analyses and surface models which are 2-D and with only some forecast time steps included. Because of all of that the list of parameters what would be beneficial to archive for UERRA is comprehensive and the way how to do it is challenging. The list contains currently more than 60 surface (single level) parameters, 12 parameters on model levels and 11 parameters on pressure and height above surface levels. In total it is around 100 parameters. In following period it will be crucial to find the best compromise between this preliminary requirements and technical feasibility what can be achieved in due course of the project with available resources.

The work on UERRA parameters' list will be done also in cooperation with other groups at ECMWF like ERA-CLIM team (producing world leading global reanalyses) or the just being created new COPERNICUS department which has recently been selected to provide climate change services for European Commission [5]. The aim is to have the same strategy in understanding of current and future users' needs regarding meteorological reanalyses and how to fulfil them.

1.5 Data handling tools

There are several tools necessary for smooth processing and maintenance of EURO4M/UERRA datasets listed below. Those tools are based on already existing technologies used for other TIGGE-like datasets in the past. Obviously many modifications have been necessary to make those tools fully compatible with EURO4M/UERRA data.

List of the data tools:

- **Data processing and archiving suite**
  Collection of Shell / C / python codes for data processing (retrieval, encoding, modifications, verification etc.) run under ECFLOW (ECMWF task monitor scheduler)
- **Data portal**
  Web interface to MARS for easy data access
- **ECMWF WEB-API [6]**
  Batch data access to data archived in MARS
- **METVIEW [7]**
  Visualization and data handling software (see Section 2 for more details)
- **ECMWF GRIB-API [8]**
  Application program interface accessible from C, Fortran and Python programs developed for encoding and decoding WMO GRIB edition 1 and edition 2 messages

The ECMWF WEB-API, METVIEW and GRIB-API are standard tools provided and maintained by the Centre.

The data portals for UERRA will be analogical to the one shown in Figure 2 and all related infrastructure is ready waiting just for new datasets to arrive.

1.6 ODB archiving

The UERRA gridded datasets on model's domains described in the previous chapters represent hundreds of terabytes of data and the related work on their archiving will consume the major part of resources of the WP4. Apart of it there is another important type of data, rescued observations by the WP1, required to be archived in MARS. The sizing of that data is in megabytes and the chosen format is ECMWF's ODB (Observation Database Format) [9].

The cooperation with WP1 has started. The first rescued observations have been coded into ODB by WP1. During the data checking by ECMWF's some issues were found in the resulting converted files. To achieve MARS compliant ODB format the WP1 group was provided with the document prepared by ECMWF's ERA-CLIM group how to code the data a better way. Also a correct ODB data sample was coded based on the original WP1 data and the coding script was provided for further reference.

Section 2: Data visualization

METVIEW is a powerful data plotting and handling software package developed at ECMWF. METVIEW is a meteorological workstation application designed to be a complete working environment for both the operational and research meteorologist. Its capabilities include powerful data access, processing and visualisation.

Any field from UERRA datasets can be visualized by METVIEW and the plots can be highly customized. METVIEW is an open source code project so any user can install and use it in accordance with open-source licence. Another way how to get graphic products (graphs, epsgrams, maps etc.) from data available at ECMWF is via implemented service OGC WMS (Open Geospatial Consortium Web Mapping Services). There are defined products which can be retrieved on demand by external user via this service. Some sample UK Met Office data from EURO4M dataset is already accessible via OGC WMS (Figure 3). The dataset shown is stored at ECMWF and visualized using the ADAGUC WMS at KNMI (euro4mvis.knmi.nl). This visualization service has been developed as a pilot in EURO4M and will be further developed in UERRA.

Apart from web data portals MARS is accessible via called ECMWF's WEB-API interface. MARS uses its own powerful internal language allowing all necessary data manipulations and this language is implemented in the WEB-API. At the moment similarly powerful language cannot be used in WMS service. The reason is that the related actually valid OGC standards are limited and cannot describe all specific and reach meteorological data features (e.g. reference or validation time, ensemble dimension). The longer-term aim is to officially enforce addition of those missing features into the existing OGC standards so that OGC compliant services could better use of data archived in MARS capabilities in the future. In the next phases the work will be aimed to further development and scaling up the current infrastructure (WEB-API and WMS services) to meet the UERRA requirements in terms of access to the data but also INSPIRE-compliance (INfrastructure for SPatial InfoRmation in Europe) what is another target of the WP4.

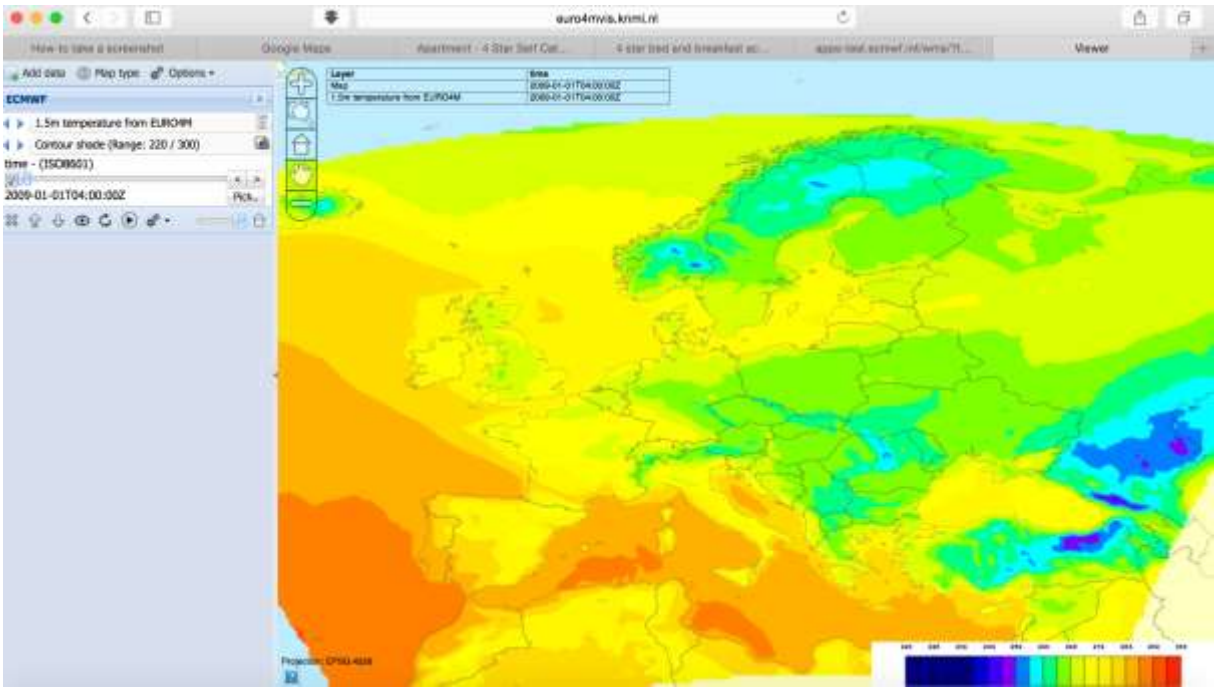

Figure 3. A sample plot produced by KNMI's visualisation tool connected to ECMWF via WMS server and using EURO4M data (UK Met Office outputs, 1.5m temperature

Annex A. References

[1] http://tigge-portal.ecmwf.int

[2] http://www.geowow.eu

[3] http://wwww.s2sprediction.net

[4] https://drive.google.com/folderview?id=0B0R5-gz4Xjz4MzJJaFBEbGhLUGM&usp=sharing_eid&invite=CI7L_tUC

[5] http://www.ecmwf.int/en/about/what-we-do/copernicus

[6] https://software.ecmwf.int/wiki/display/WEBAPI/ECMWF+Web+API+Home

[7] https://software.ecmwf.int/wiki/display/METV/Metview

[8] https://software.ecmwf.int/wiki/display/GRIB/Home

[9] https://software.ecmwf.int/wiki/display/ODB/ODB+Home

Acronyms and abbreviations

CANARI	Code d'Analyse Necessaire a ARPEGE pour ses Rejets et son initialisation (Météo-France operational surface analysis)
COSMO	The Consortium for Small-scale Modeling
DWD	Deutscher Wetterdienst
ECMWF	European Centre for Medium-Range Weather Forecasts
EURO4M	The European Reanalysis and Observations For Monitoring
EPS	Ensemble Prediction System
GCI	GEOSS Common Infrastructure
GEO	Global Earth Observation
GEOSS	Global Earth Observation System of Systems
GEOWOW	GEOSS Interoperability for Weather, Ocean and Water
GRIB	Gridded Binary
INSPIRE	INfrastructure for SPatial InfoRmation in Europe
HARMONIE	**H**irlam **A**ladin **R**esearch on **M**eso-scale **O**perational **N**WP **i**n **E**uromed
MARS	Meteorological Archive and Retrieval System
MESAN	SMHIs Mesoscale Analysis System
MESCAN	"blending between MESAN and CANARI"
MF	Météo-France
MIUB	Meteorologisches Institut der Universität Bonn
MO	Met Office (UK)
NetCDF	Network Common Data Form
ODB	Observation Data Base
OGC	Open Geospatial Consortium
S2S	Sub-seasonal to Seasonal Prediction Project
SMHI	Swedish Meteorological and Hydrological Institute
THORPEX	The Observing System Research and Predictability Experiment
TIGGE	The International Grand Global Ensemble
TIGGE-LAM	TIGGE-Limited Area Model
UERRA	Uncertainties in Ensembles of Regional Re-analysis
WMO	World Meteorological Organization
WMS	Web Mapping Services