UERRA Newsletter No 2

April 2016

Introduction

UERRA (www.uerra.eu) is a bit more than halfway through its 4 years. During the first two years the observation rescue reanalysisis, data services and quality evaluation systems have been developed. Some 8.5 M data have been processed and additional ones will be extracted from already digitised national archives. The reanalyses in 4 different versions or applications are ready to start or have started running. Now with data services and archiving started and evaluation of uncertainties beginning, everything in UERRA is falling into place and being put together to show a pre-operational Copernicus regional reanalysis service.
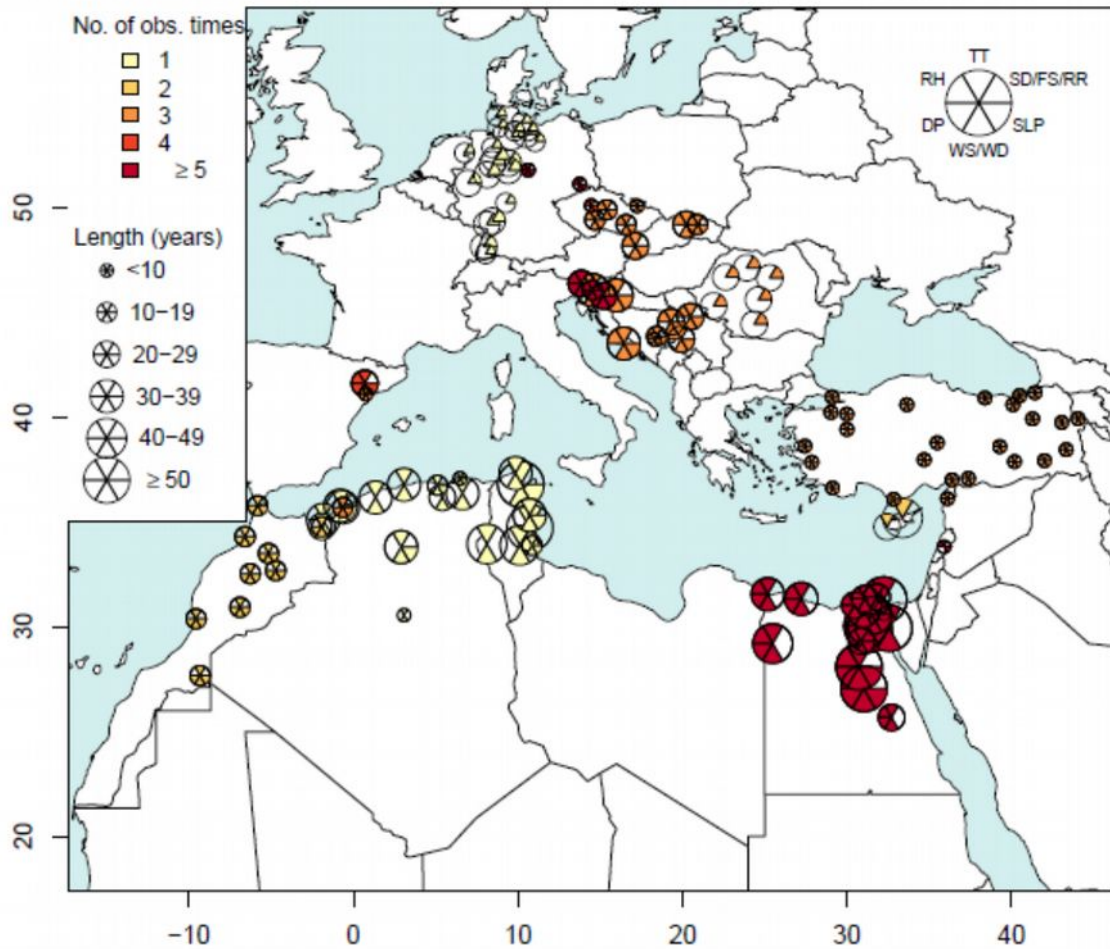
Meanwhile, the third General Assembly was held in Toulouse this February. In connection with this, there was an early User's Workshop where a good number of Users of scientists working with users, presented their applications and user requirements.

UERRA General Assembly , 02-04 February 2016
Centre International de Conférences - Météo-France - Toulouse
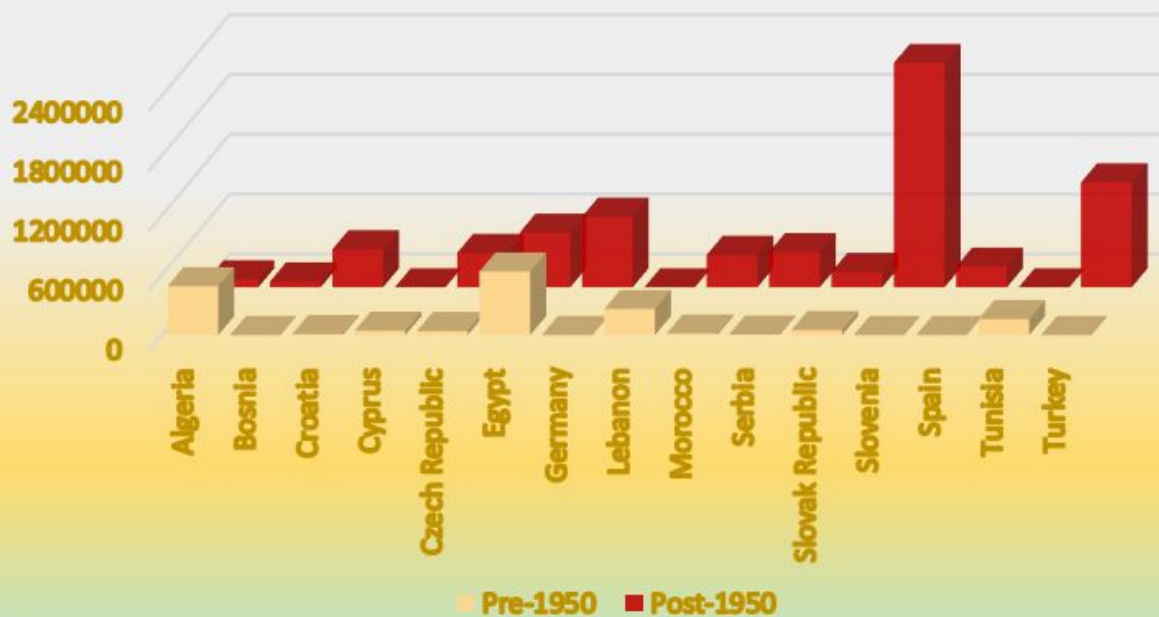
Observation data rescue and data development (quality control)

Work Package 1 in UERRA is involved in data rescue of historical observations of temperatures, precipitation but also other variables like pressure and winds. These are often observation journals on paper but may also be scanned or in some cases already digitised in national data bases that have not been transferred to international data bases yet. An Appendix (I) lists the different data bases which are targeted in UERRA and through the work of the partner leading the effort, URV in Tarragona, Catalonia, Spain. WP1 is involved also in many data rescue efforts like ACRE, ERA-CLIM2. I UERRA the efforts have concentrated on post 1950 data and on observations during different times of the day as well (in addition to daily values. Data are from several European countries but far from all due to availability or data polices.

UERRA data sources 1879–2012

Data volumes by country

The digitisation has progressed faster than anticipated due to more experienced staff guiding and supervising students doing the work at URV. There have been over 8.4 M data processed from about 100 stations in several European countries and some North African Mediterranean ones and Turkey. The emphasis is on post-1950 and on observations on the sub-daily time scale (often 3-5 observations per day). This has required improved quality control procedures and they have been developed. URV has carried out almost all of these whereas NMA-Romania has contributed with precipitation data from some Romanian stations. Extensive quality control has been applied both in manual and automatic ways. With feedback and supervision the error rate has been pushed down to very few % of the data. Homogenisation is then the final step before data can be disseminated. Some of the data can also be used in the ongoing reanalysis efforts in WP2 whereas others will benefit all future reanalyis efforts. In addition, already digitised and quality controlled data from Sweden and Norway have and are being collected for combing into the data bases.

Reanalysis systems and ensembles

Work package 1 also deals with interpolation of processed observations on the European scale and connects and interacts with European efforts on the global as well as regional and national scale. In particular the interpolation methods are being developed to take better account of the physical structures of precipitation and temperature and with respect to the land orographic structures. In addition, as it is in UERRA, uncertainty estimation evaluation methods have been developed.
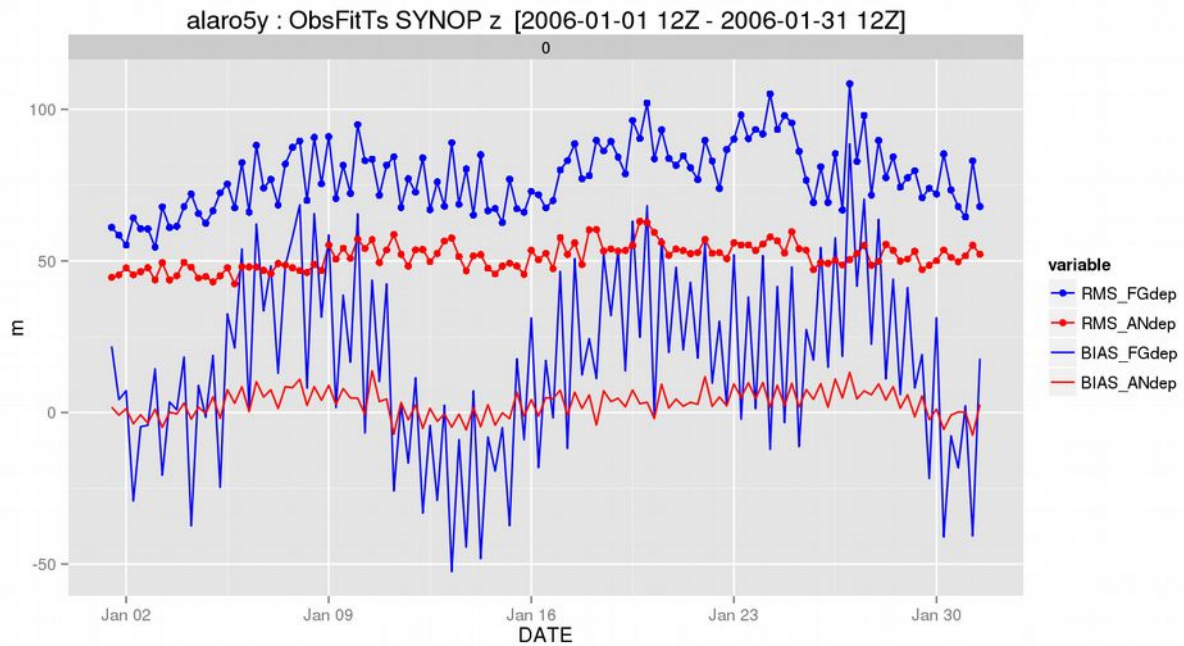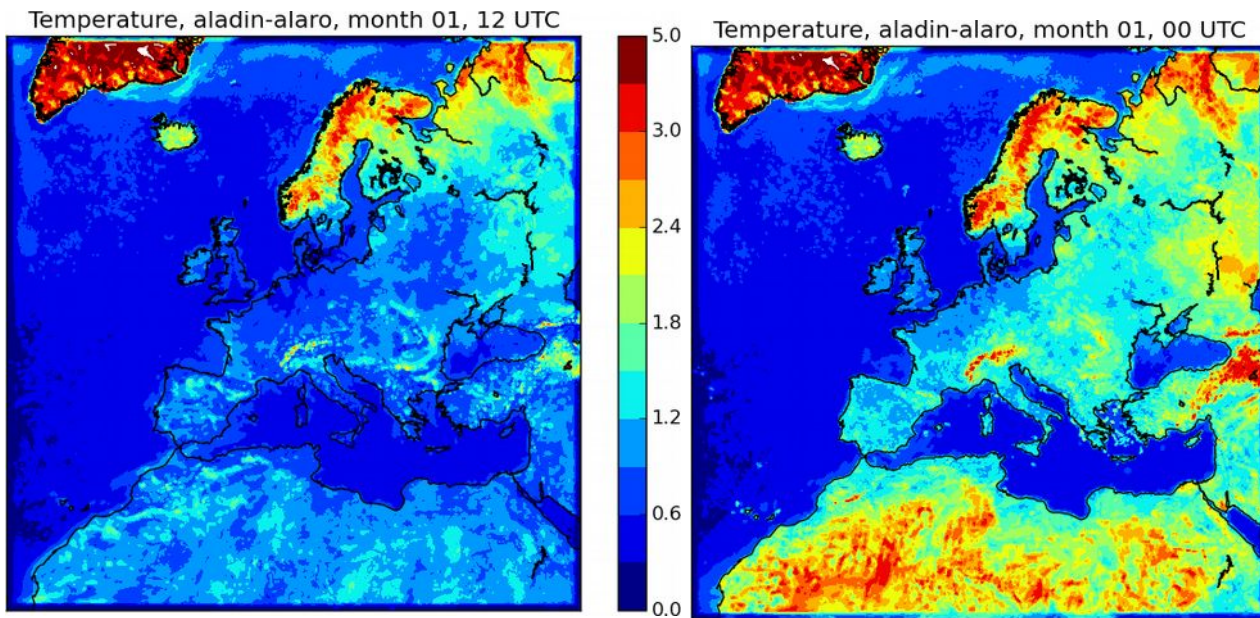
More data and new versions of data sets (E-OBS and CRU in particular) have been released. There is a much increased interest in very recent data analyses in near real time due to climate variations and recent records seen in the climate change.

In UERRA the gridded observational data sets are seen as complimentary information that may even be used for some data that are not available otherwise. Otherwise, the main connection with the main reanalysis Numerical Weather Prediction (NWP) model based reanalyses, is for validation and use as a mainly independent data set. The Work Package 3 in UERRA works with such methods and with estimation of uncertainties in the reanalysis and will use both these moderate resolution but all area covering, data sets and supplemented by regional and national data sets produced with other resources but provided to UERRA through the UERRA partners and scientists. The more local data sets are of very high resolution and allow for detailed analysis of error structures and particularly of correlations both in space at high (2 km) resolution and in time.

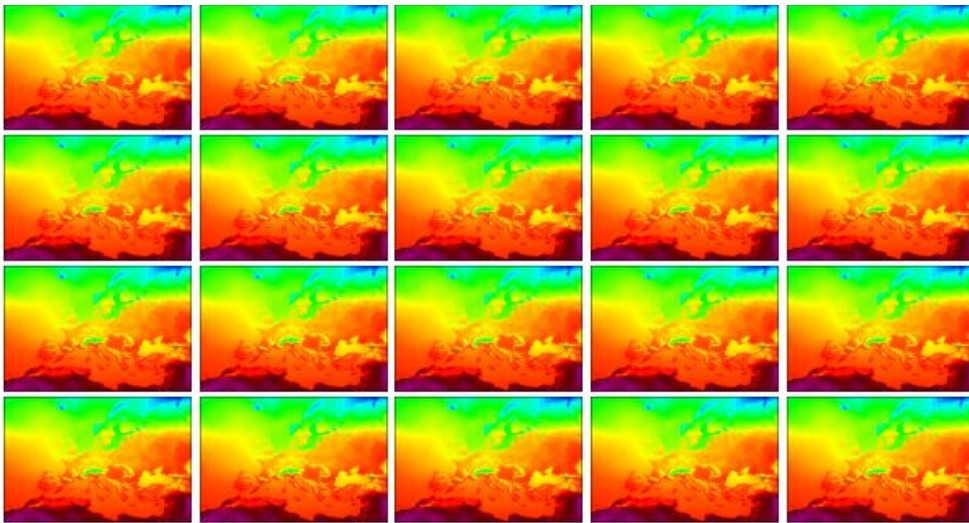# NWP based reanalyses developed in 4 different flavours

There are three different fully 3-dimensional reanalysis systems developed in UERRA and a fourth one which is analysing near surface parameters only but at higher resolution, 5 km, instead of about 11 for the three upper air reanalyses.

SMHI has run two 5-year reanalyses (2006-2010) with two model versions, both to provide some feeling for the uncertainty or differences due to model differences but then in order to chose the best version for the full 55 years of reanalyses. Some examples of the differences and standard deviation of differences are shown below:

Temperature, aladin-alaro, month 01, 12 UTC — Temperature, aladin-alaro, month 01, 00 UTC

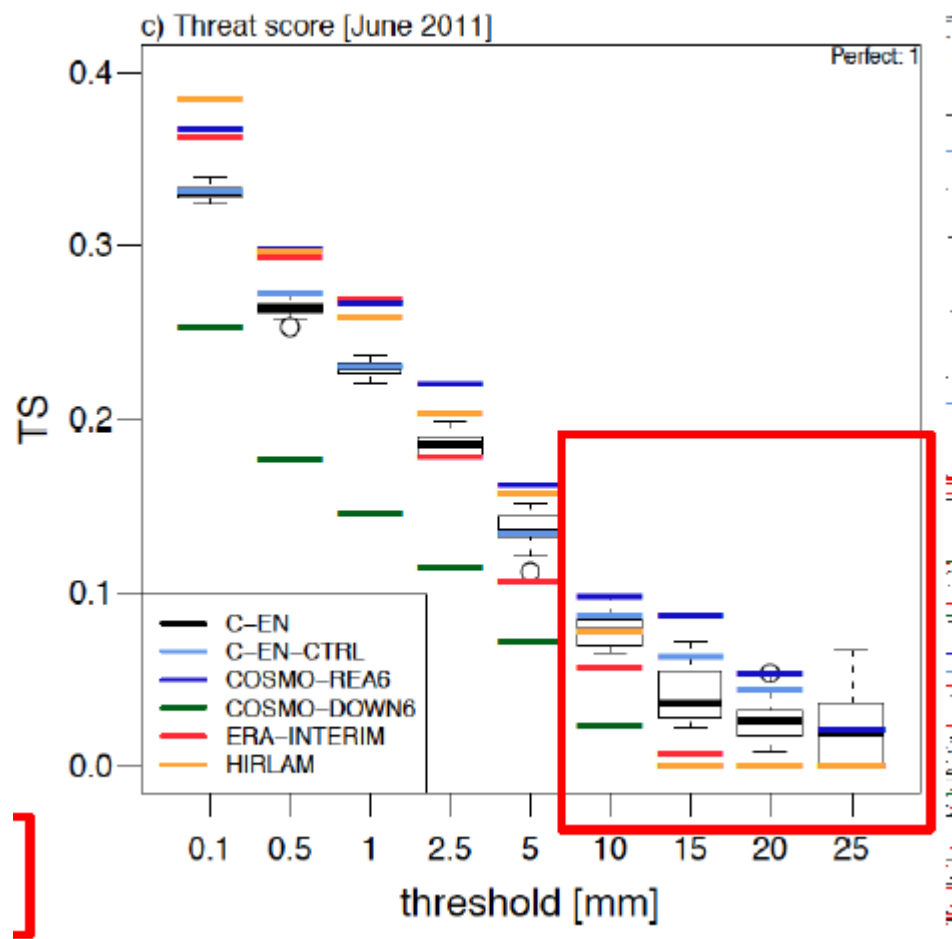alaro5y : ObsFitTs SYNOP z [2006-01-01 12Z - 2006-01-31 12Z]

The use of observations and the fits to the short range forecasts (6 hours) used in the reanalyses and the reanalyses themselves have been monitored regularly through the 5 years. An example fro January 2006 is shown above.

The Met Office has developed an ensemble assimilation system with 20 members of 4D Variational Data assimilation (4D-Var DA) but at 24 kilometre resolution. This has been run for test periods for about a month and is ready for production. The ensembles provide some estimate of uncertainties through their spread and furthermore, the statistics will be fed in to the deterministic 12 km 4D-VAR to give a flow (or weather situation dependence). This whole system will be run for the satellite era, from 1978 for at least 25 years. The spread in the ensembles can be seen in an example below.
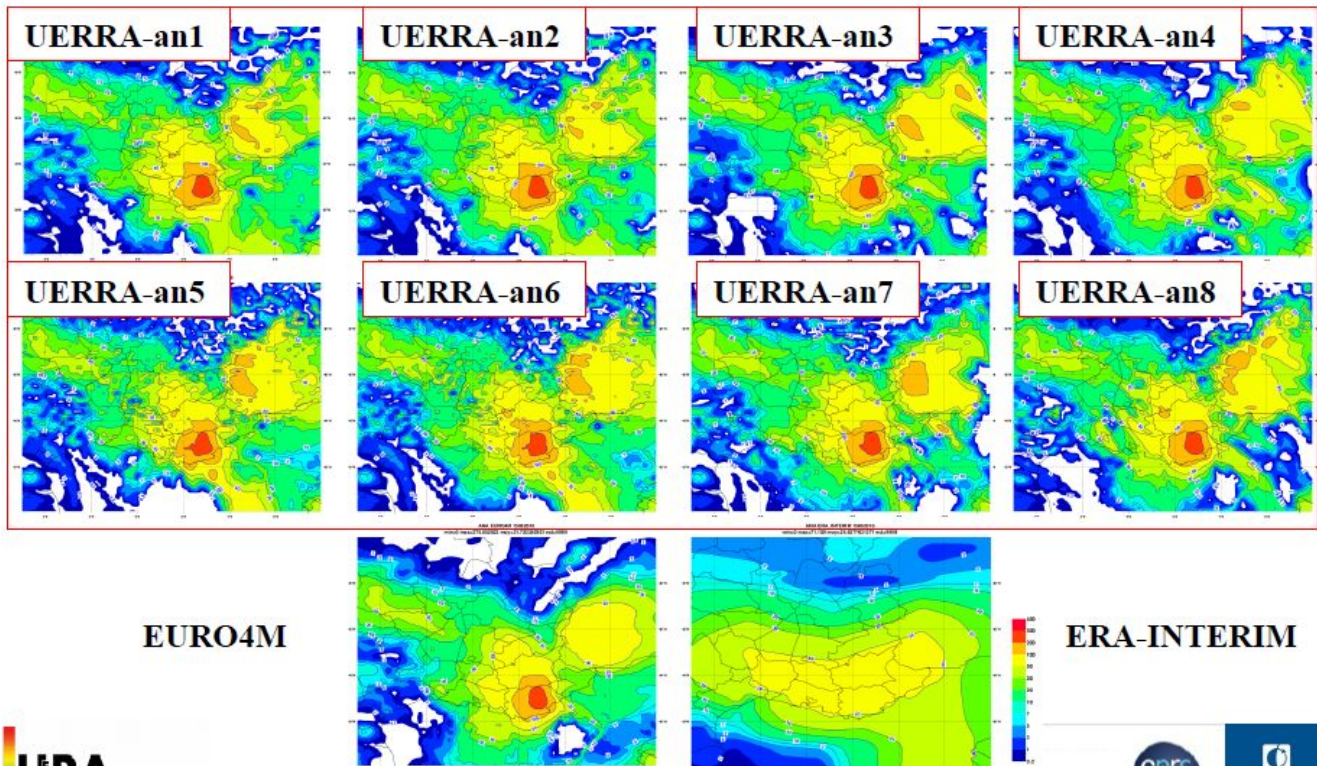

The University of Bonn has developed an ensemble assimilation with the COSMO model at 12 km horisontal resolution over the same Euro-Atlantic region as the others. Although an more advanced ensemble system is under development, the nudging based ensemble is starting to be run now (spring 2016) with 20 members. The demonstration of the method for reanalysis and ensemble spread based uncertainties will be limited to 5 years of assimilation.

c) Threat score [June 2011]

Perfect: 1

Legend:
- C–EN
- C–EN–CTRL
- COSMO–REA6
- COSMO–DOWN6
- ERA–INTERIM
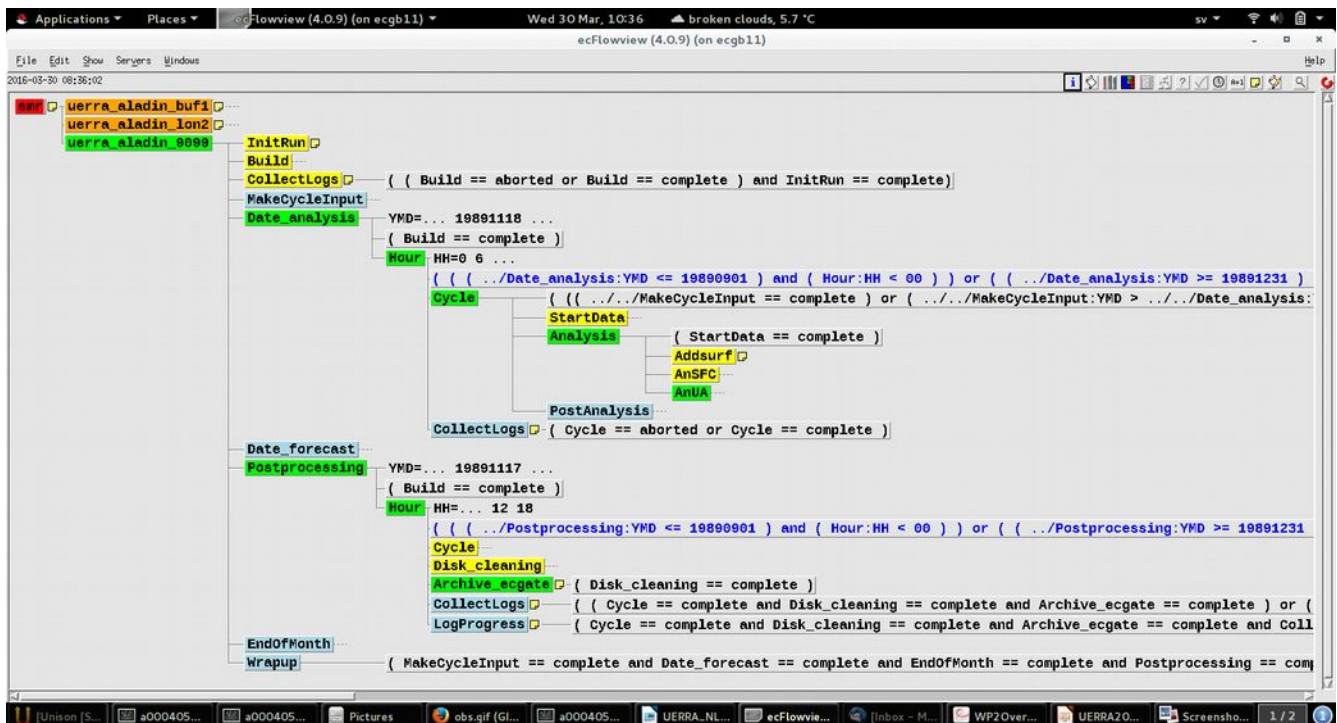- HIRLAM

Y-axis: TS
X-axis: threshold [mm]

Météo-France has developed a near-surface parameter ensemble reanalysis system (for precipitation and temperature and humidity at 2 metres in the UERRA application). It is using results from the SMHI upper air reanalysis with the two model physics versions at 11 km resolution and then downscaling them with different methods and model versions. This will give a sort of ensemble with at least 6 members to be run for 5 years (2006-2010) at as high resolution as 5 km. A long, 55 years, run at 5 km will also be made, from and like the SMHI reanalysis at explained above.
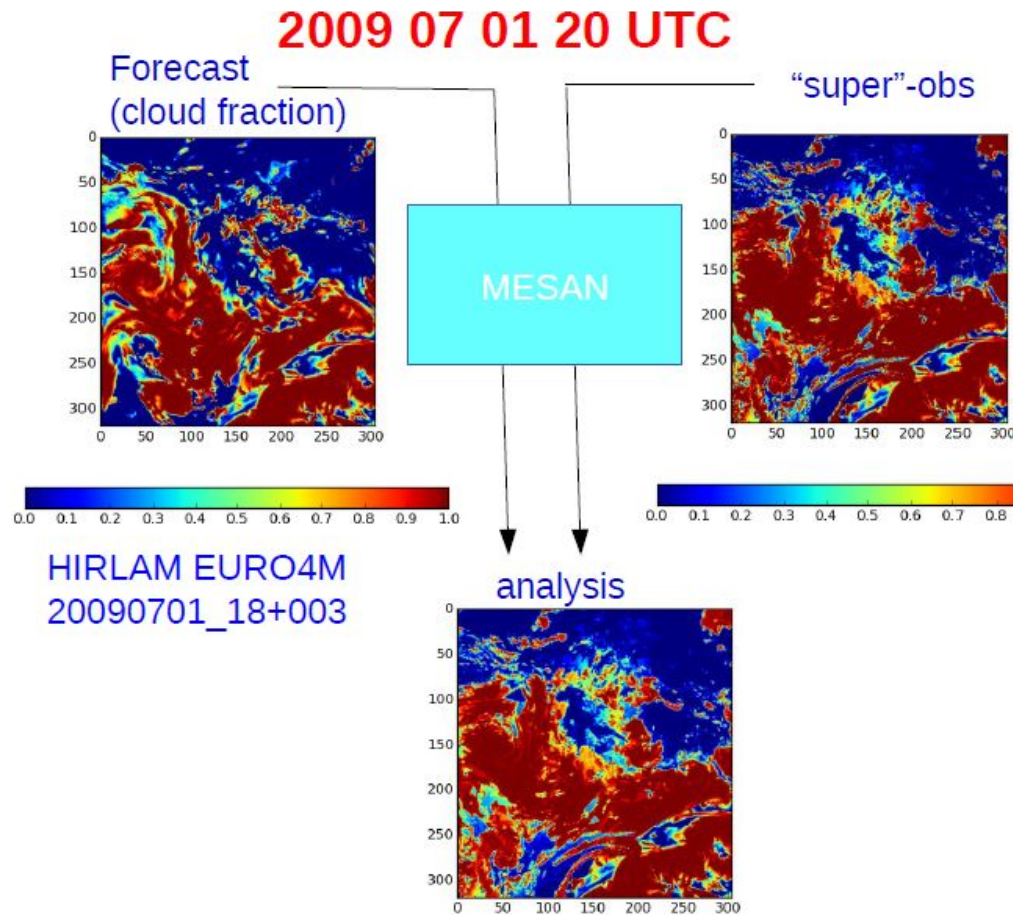
Extreme precipitation events of 15 June 2010
8 members : RR24h UERRA Analysis

UERRA-an1 · UERRA-an2 · UERRA-an3 · UERRA-an4 · UERRA-an5 · UERRA-an6 · UERRA-an7 · UERRA-an8

EURO4M · ERA-INTERIM

GA3 UERRA / Toulouse, 1-3 February 2016

Rather complex systems of scripts, jobs and scheduler instructions have been set up at all the participating centres. An example of the control in one reanalysis cycle for SMHI at ECMWF is shown below. All the applications of the analysis and model systems have been set up in unique domains, resolutions and parameter settings for this European-Atlantic UERRA area and they are different from any of their operational runs and settings.

Another and 5th special reanalysis for 25 years of satellite cloudiness data has been developed at SMHI. This will mainly be based on the geostationary SEVIRI and MVERI instruments but also on cloud mask data from polar satellites and also using a model background from the HIRLAM data in the precursor EURO4M project. It will use the latest and best available EUMETSAT SAF products. The results of the cloudiness reanalysis will also be used in the evaluation part of UERRA.

2009 07 01 20 UTC

Forecast (cloud fraction)

"super"-obs
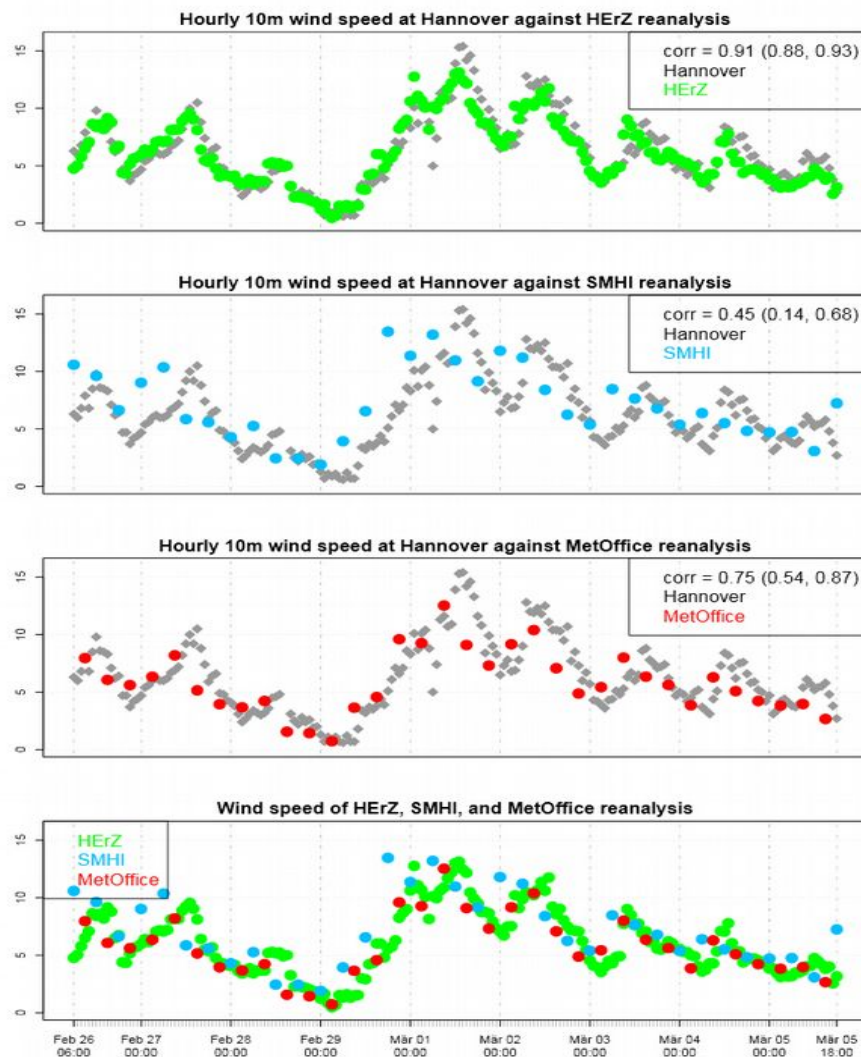
MESAN

HIRLAM EURO4M
20090701_18+003

analysis

Uncertainty and quality evaluation and validation

A comprehensive software package for evaluating reanalyses in several different ways has been further developed. The package has been built at DWD and MET Norway and based and evolved from previous packages in R and HARP from the HIRLAM consoritum. A github repository has been set up where partners can contribute and the software can be shared and it is openly available.

Since the UERRA production of reanalyses has just started (depending on which of the 4 partners you are interested in; there is data from some years from SMHI but not the others), the evaluations and tests
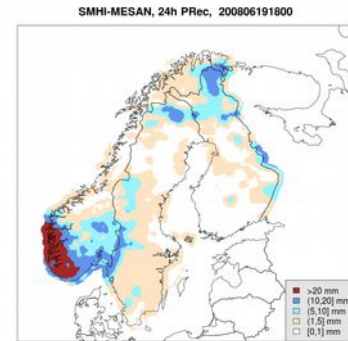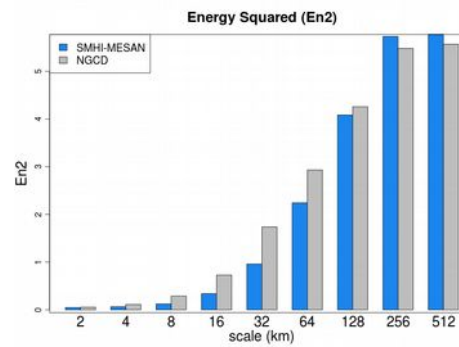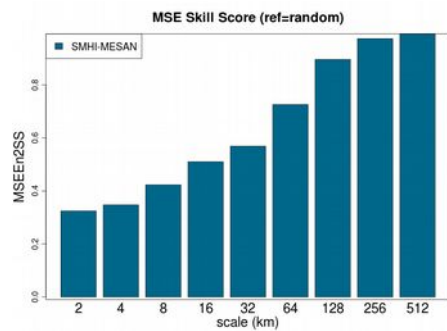
have been done on the earlier EURO4M Project reanalyses. 2 years of such data
have been archived in MARS at ECMWF. In addition there are COSMO reanalyses at DWD (Hans
Ertel Centre for Weather research) that have been used as input, both to evaluare results and
uncertaintis, but also to test the data access and the funtions of the software.

The different reanalyses have been intercompared and compared against observations, both
conventional and then independent ones (not used in any reanlysis) from mast data. Below time series
against station data are shown for a few different reanalyses (note that EURO4M reanalyses are
different from UERRA, in terms of models, methods and resolution. The COSMO HerZ is at high
resolution (6 km). The comparisons have started with conventional scores (correlation, bias, RMSE,
anomalies, frequency distribution, contingency 4x4 yes/no table for hit rate and false alarm etc.). More
have been added like PDF (probability density function) based scores. Recently satellite data have been
included in the evealuation and in terms of radiances which are measured and computed from the
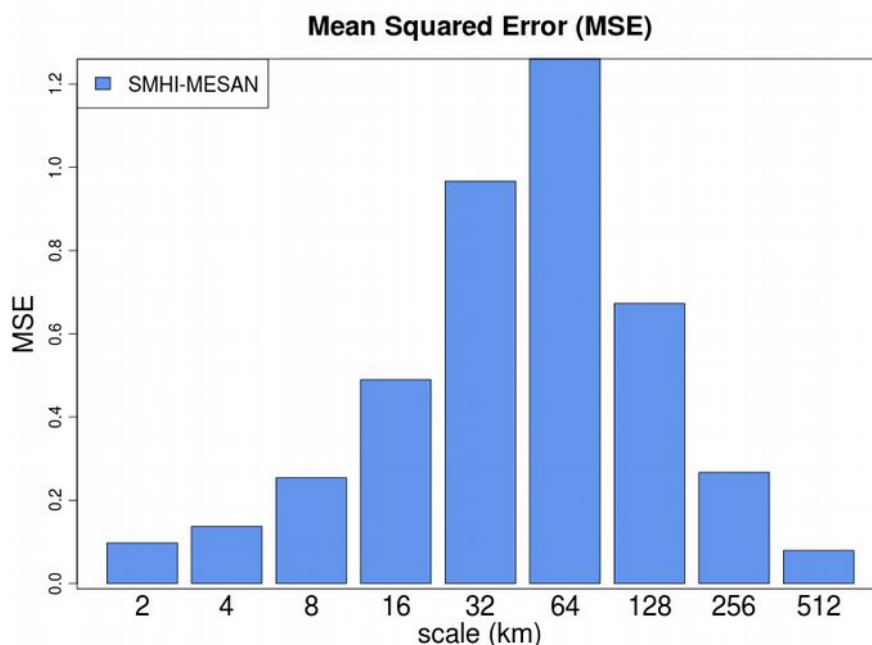models in the reanalyses.

Another important way to test the quality is to user high resolution observation based reanalyses done
over subregions or nations in Europe. Such ones are over the Alps (Meteo-Swiss) and over the Feno-
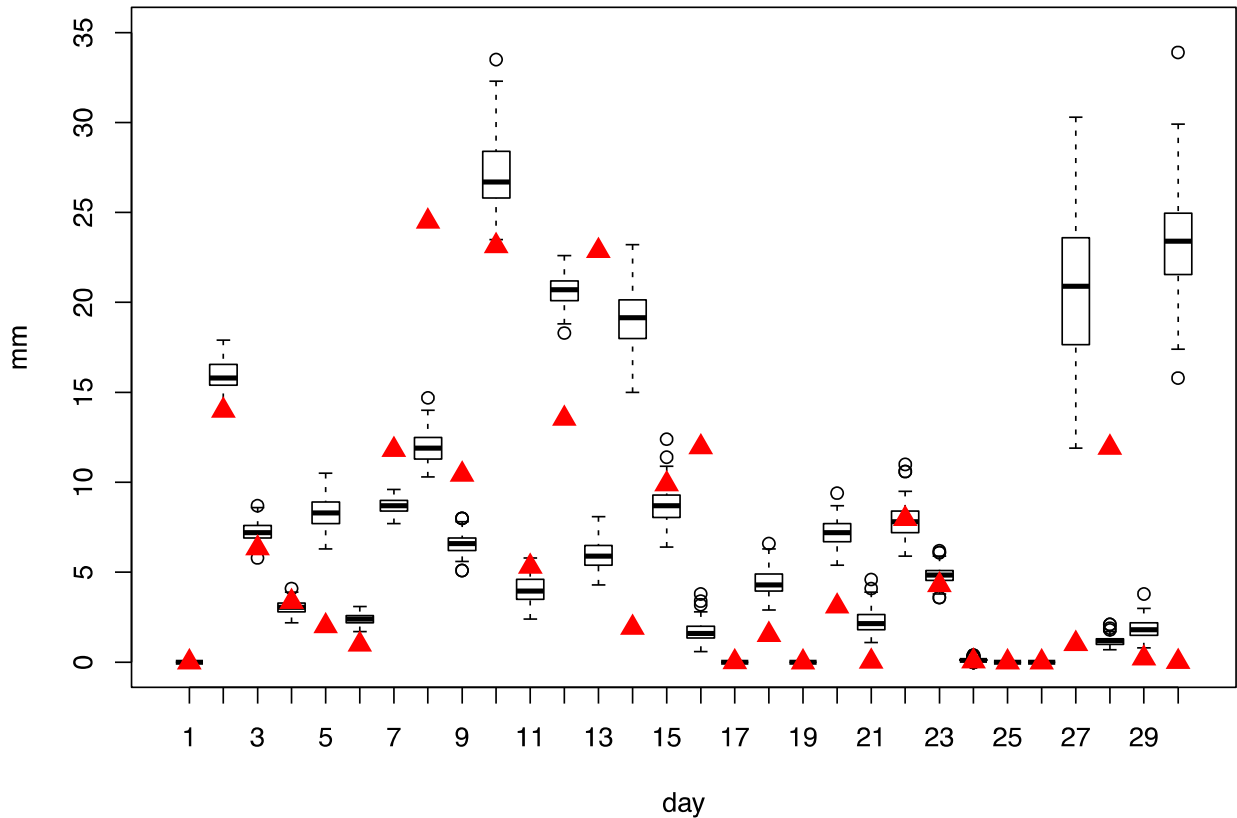Scandian domain (MET Norway).

The advantage of the high resolution data is to verify at different scales separately. One can see that is more skill due to higher energy in the larger scales in the SMHI MESAN analysis (a few 100 kilometers). Another method that has been employed is and intensity based technique where one separates according to thresholds and also scales.

The observation based data sets at KNMI and Meteo-Swiss have now been complemented with probabilistic estimates of interpolation errors and uncertainty bars can be put on the data. An example is shown below where the old HIRLAM EURO4M data are compared with the probabilistic high resolution interpolation. It can be seen that for some days the HIRLAM data are outside or even far outside the range of uncertainty and thus almost certainly in gross error.

A296 199006

Data services and visualisation

First a between-centres common two year EURO4M reanalysis period was archived in MARS at ECMWF as a pilot data set and for the evaluation team to work with. These data are however quite heterogeneous in their choice of parameters and types and times of data (e.g. analysis or forecast and surface or upper air) since there was no prior coordination at this level of detail.

In UERRA there is a common set or parameters (for most of them; there are limitations or extension in the different models that result in that a few parameters are not in all models or reanalysis producer's archives). They have been selected and iterated between the partners and taking user surveys and interactions into account. Moreover, they will all be converted to GRIB2 and conform with WMO standards (and extending the standard where necessary).

It means also that a lot of tools for data services like webapi and metview at ECMWF can be applied to the UERRA MARS data.

ECMWF

About Forecasts Computing Research Learning

Navigation

Job list
MARS activity

See also...

FAQ
Accessing forecasts
GRIB decoder

MARS Catalogue

Date (1 values)	Time (1 values)	Parameter (1 values)
2010-01-01	00:00:00	Temperature

▶ Check for availability
▶ View the MARS request
▶ Estimate download size
▶ Retrieve the selection in GRIB or NetCDF (experimental)

Note about availability

Some of the fields may not be archived at all levels or all forecast time steps. Before retrieving data you may want to check the availability of the requested fields. For that, follow the *Check for availability* link.
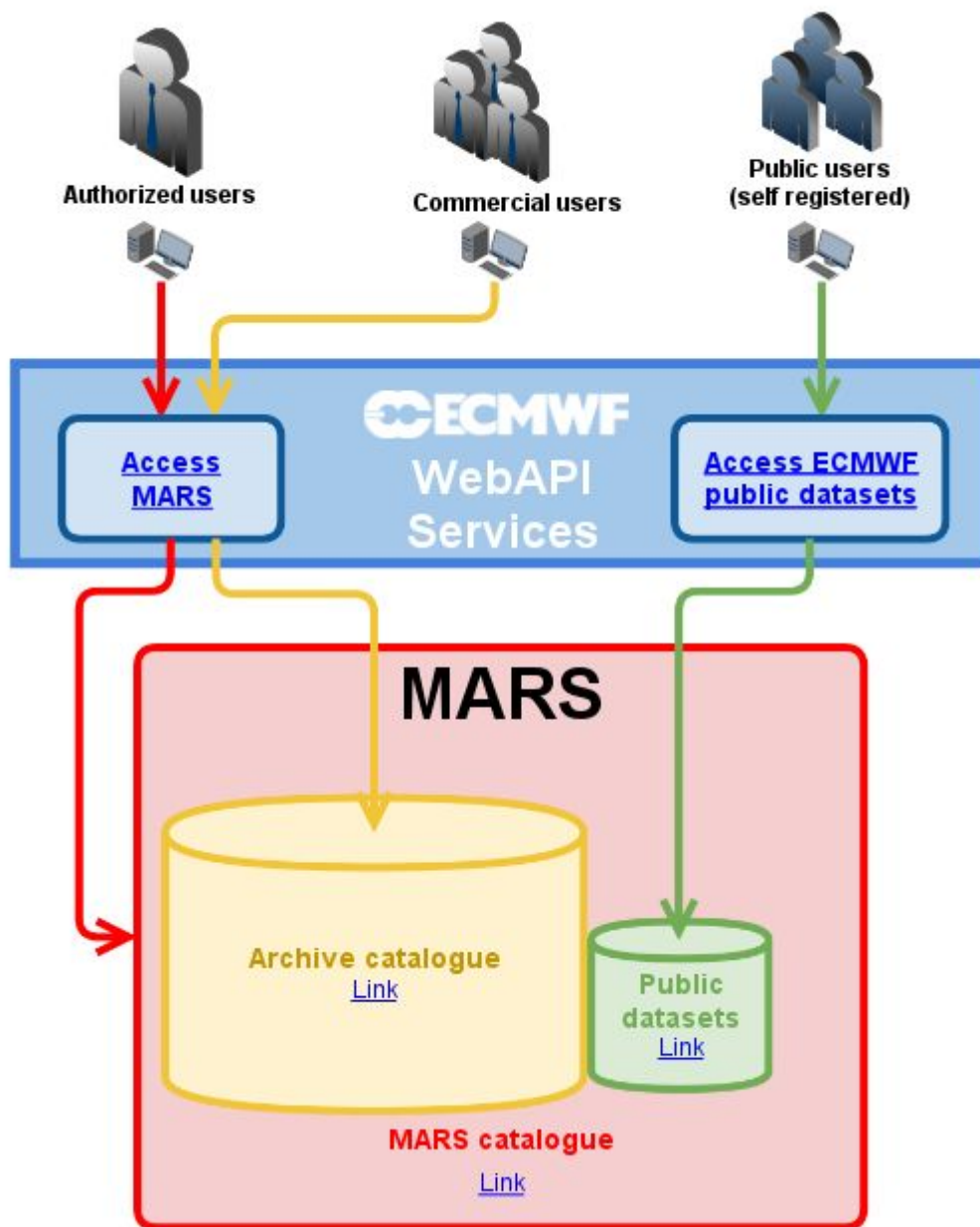
Retrieving

In order to retrieve data, you must select at least one item in the lists above. You can select more than one item in each list.
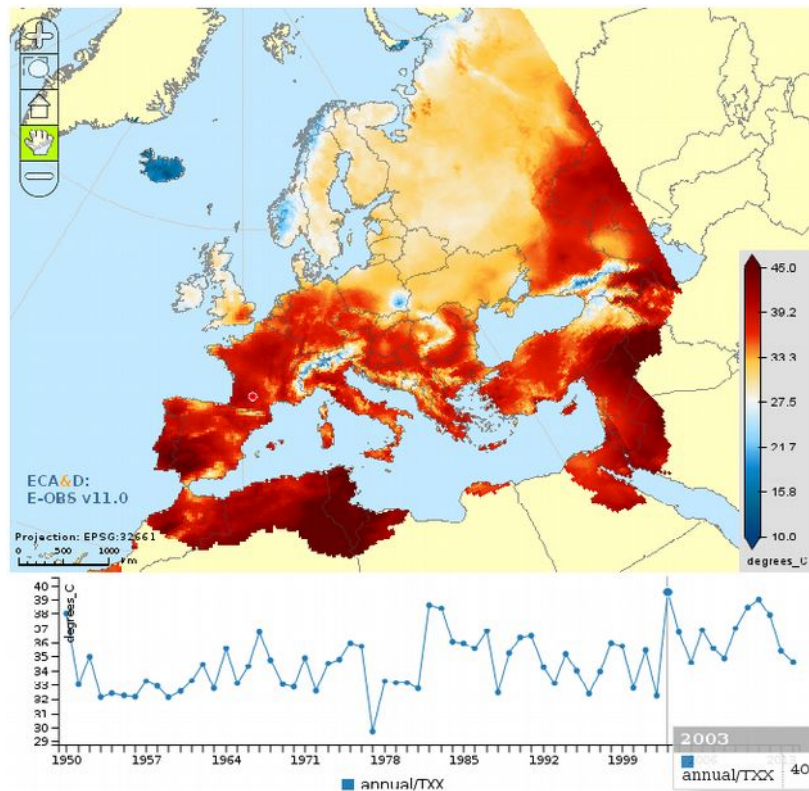
Current selection:

en . ep . er . gr . j5 . la . mc . ms . nr . od . rd . rm . s2 . se . sr . ti . tr . uk . ur

▲ Top of page

copyright © ECMWF

Visualisation and data services can also be applied on some of data through the tools at KNMI.The ADAGUC tools at KNMI can be applied on many different data sets including the EURO4M and
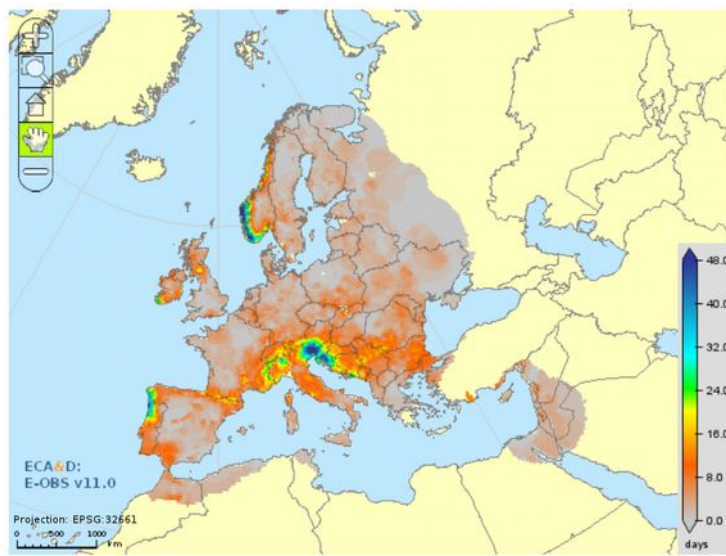
UERRA ones.

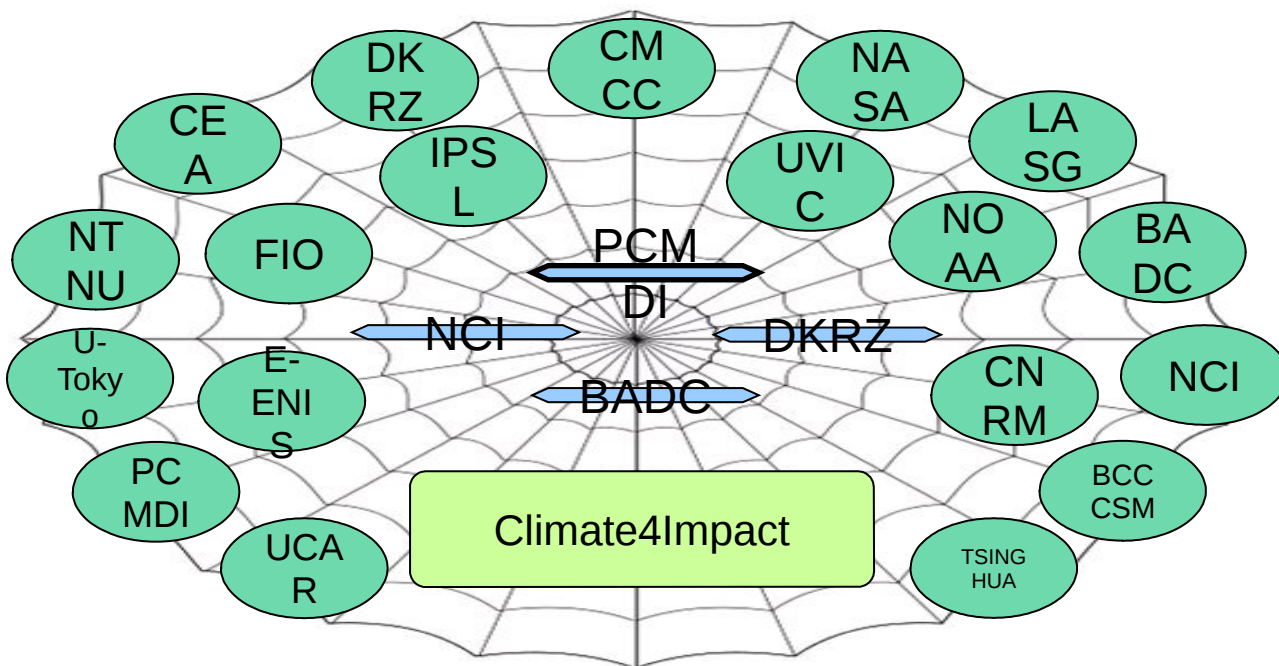## User oriented products, contacts, outreach and dissemination

There are users requesting EURO4M data sets archived in Sweden (SMHI and NSC) and some of it through the CLIPC project, others though voluntary help from SMHI. There is a huge potential due to the long 25-35 time series present in some of the data sets.

User oritented products continue to be constructed from the E-OBS data sets, both from observations but mainly from the interpolated, gridded data sets. Climate information bulletins are regularly published and below is shown the number of day with heavy precipitation.

The ADAGUC Web map server is available at KNMI since 2009 and can be used to display and overlay all sorts of data inlcuding observations, from NetCDF files.

ECA&D:
E-OBS v11.0

Projection: EPSG:32661

ESGF (Earth System Grid Federation) is used for distributed data in the climate community and will be demonstrated from some of the UERRA data as well. This is linked with the work in the adjoining FP7 CLIPC project where a Climate Information Portal is the core of the project.

There is also a part of the project that deals with hydrological modelling where the input comes from the reanalyses. A first such work was done with the EURO4M data as input and more will follow when the UERRA reanalyses are available.

A User workshop (of reanalyses) was organised in Toulouse 3-4 February 2016 and it was well attended by both UERRA data producers and evaluators (in terms of QA) and users from the climate change research and energy industry and some commercial applications like insurance and media.

Some 20 users described and showed their current and projected needs of climate reanalysis data. Many want it at very high resolution both in space and time and over long multi-decadal time extent. Also a dense vertical resolution in the lowest part of the atmosphere is needed. The following is a condensed version of what came out from the Workshop and it quite typical of user needs but not exhaustive as certainly not all types of users could be represented here.

- Variables:
  - 2m temperature, solar, direct and global radiation, wind
  - wind at 50, 70, 100, 150 m and more, wind gusts
  - profile of water vapour and temperature
- Spatial resolution
  - From 25 km to 10, 5, 2.5 km and even higher, below 1 km and a few 100 m (downscaling?)
- Time resolution
  - Generally 1 hour
- Time period
  - 30 – 50 years or more
- Applications:
  - Downscaling climatology
  - Forcing high resolution models
  - Validating models
  - Energy:
    - wind potential and climatology, balance in the grid
    - Solar energy and potential and climatology
  - Climate modelling

- Climate services

- Insurance and severe weather statistics, extreme values

- Agriculture (and forestry):

  - Growing season index

  - Winter severity

ANNEX I

LIST OF GLOBAL DATA CENTRES TO WHERE THE RESCUED AND ACCESSED DATA UNDER UERRA WILL BE PROVIDED, ONCE THE OBSERVATIONS HAVE BEEN DEVELOPED (CHECKED FOR QUALITY AND HOMOGENEITY)

As described in Deliverables D1.3 and D1.4 of UERRA, all the observations recovered, digitised, quality controlled and homogenised under the UERRA Data Rescue effort by URV and NMA-RO, about 8.4M station values, in addition to those data accessed in digitised format by URV from several National Meteorological and Hydrological Services (NMHS) in Europe (Norway and Sweden, and Catalonia and Slovenia, if finally verified), will be made publicly accessible and provided to the MARS Archive at ECMWF through the following global and regional data centres (the order shown indicates a priority order):

1. The International Surface Pressure Databank (ISPD) at the National Center for Atmospheric Research (NCAR)
   All the air pressure observations digitised or accessed will be provided in ASCII format to the ISPD ensuring their provenance once quality controlled and tested for homogeneity (around January 2017) and on time for the new version (ISPDv6). The ISPD is the authorised databank from where the MARS Archive gather all the relevant observations at the synoptical scale.
2. The ZENODO scientific data repository
   All the observations recovered and accessed under UERRA will be made publicly available through this global repository for scientific data, including observations for air pressure, temperature, dew point temperature, wind speed and wind direction, relative humidity, fresh snow, snow depth and precipitation.
3. The European Climate Assessment and Dataset (ECA&D) at the KNMI, De Bilt, Holland
   All the recovered and accessed observations will be provided to ECA&D in two versions: the quality controlled version and the homogenised version.
4. Other global data centres and databanks to which all the observations recovered and accessed under UERRA will be also provided in ASCII format and the two versions produced to the following databanks and initiatives:
   a. The Met Office Hadley Centre observations datasets: the HadISDH dataset
   b. The International Surface Temperature Initiative (ISTI) databank

In addition, the data digitised, quality controlled and adjusted under UERRA thanks to the access provided to the original data sources by some European NMHS (e.g. DWD, the Slovenian and the Catalan services) will be sent to the respective NMHSs once the data have been developed. It's is expected to provide the data recovered and accessed in their different versions to the aforementioned data centres and NMHSs not later than two months after the completion of the D1.6 foreseen for month 36 (M36) of the UERRA project.