UERRA Newsletter No 3

January 2017

Contents

<u>Introduction</u>

UERRA (www.uerra.eu) is a now in its last (4th) year. All the pieces are now falling into place:

- Data rescue and development has been completed in UERRA with much more data than anticipated and data development (quality control and homogenisation) has been done on now also the sub-daily time scale;

- Improvements have been made to the E-OBS data set, in methodology and in assessing uncertainties;

- The Regional Reanalysis production is in full swing with significant parts of the time periods covered.

- Conversion of Reanalysis data to WMO GRIB-2 has been accomplished and archiving in MARS is going on as fast as possible for the common UERRA archive;

- Developments of evaluation methods and the common software package has been completed and tried on some of the Reanalysis data;

- User contacts and meetings and workshops have taken place with users and also with related Projects and EU stakeholders;

- The above components are preparing for operational Copernicus Climate Change Services for which tenders have been issued during the latter part of 2016.

Production Renanalyses in UERRA

The Met Office, SMHI, Météo-France and University of Bonn area each running their European Regional Reanalyses in production and are at varying advanced stages of production. Moreover, the ECMW MARS archives of UERRA are being populated with reanalysis grid point fields in GRIB2. Below is a current (January 2017) picture of the status and plans for the rest of the project. Certain periods like 2005-10 and 1981-85 are of priority to be completed for all (not Univ. of Bonn since only 5 years are planned for them). (Thanks to Richard Mladek, ECMWF for making and updating them on the ECMWF wiki web site).
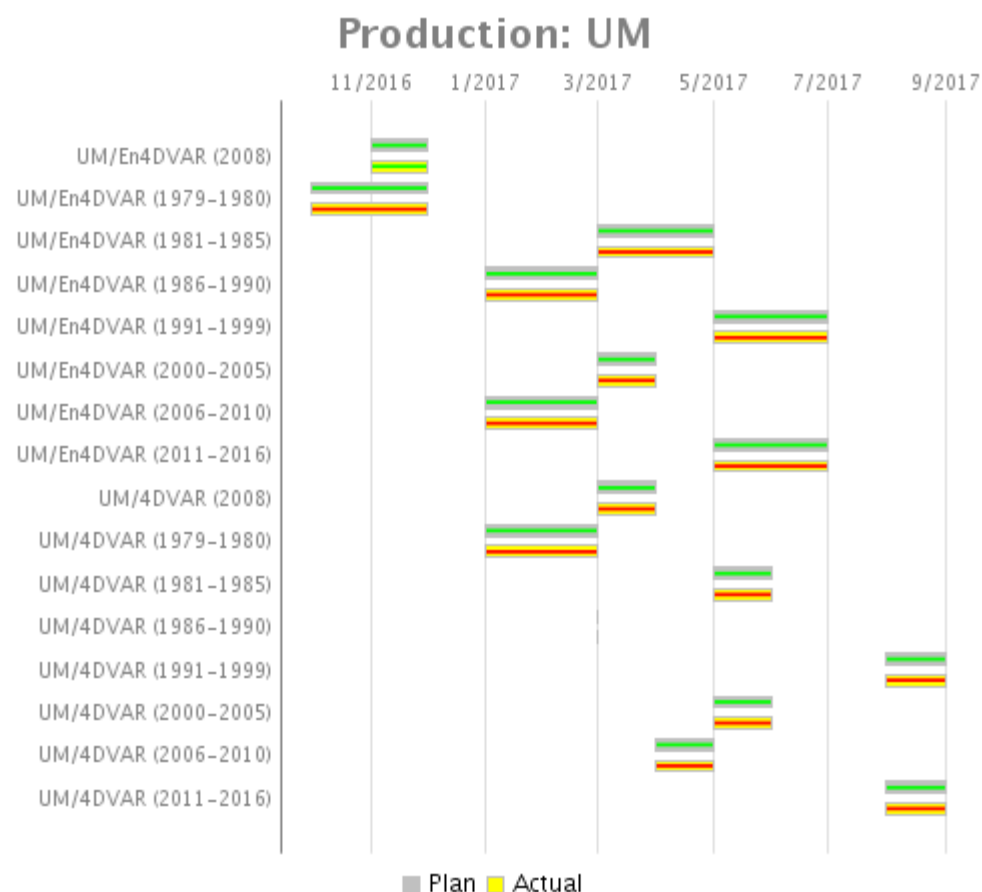

Figure 1. Met Office production streams with Unified Model (UM) regional UERRA reanalysis. Dark green is the planned, light green done and red to be done.

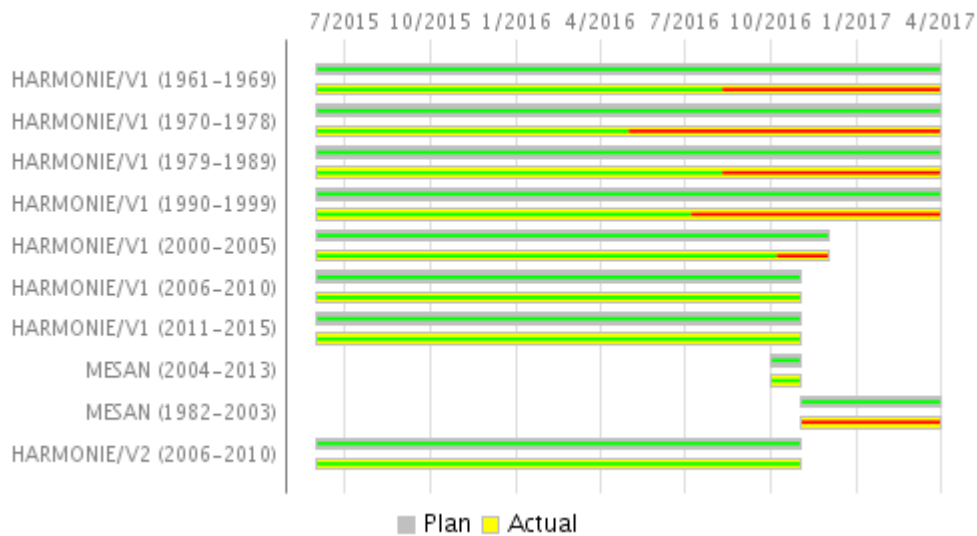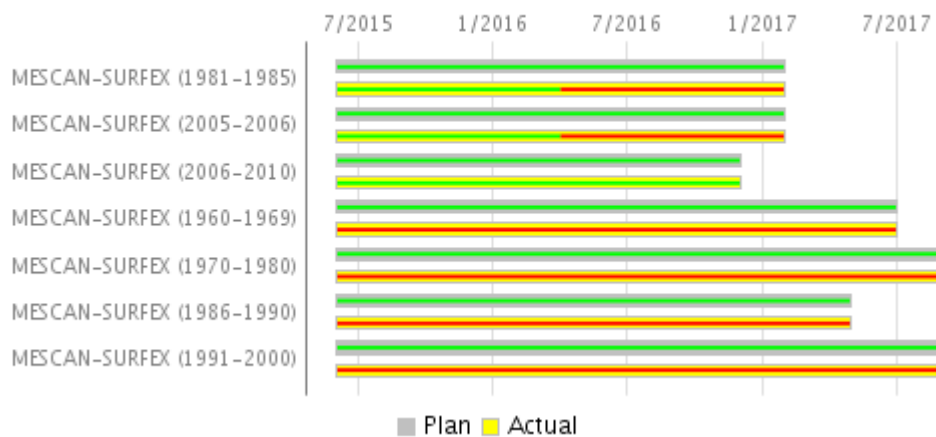*Figure 2. As Fig.1 but for SMHI HARMONIE (ALADIN and ALARO)*

Figure 3. As Fig.1 for Météo-France and MESCAN 2D analysis and SURFEX modelling from that.

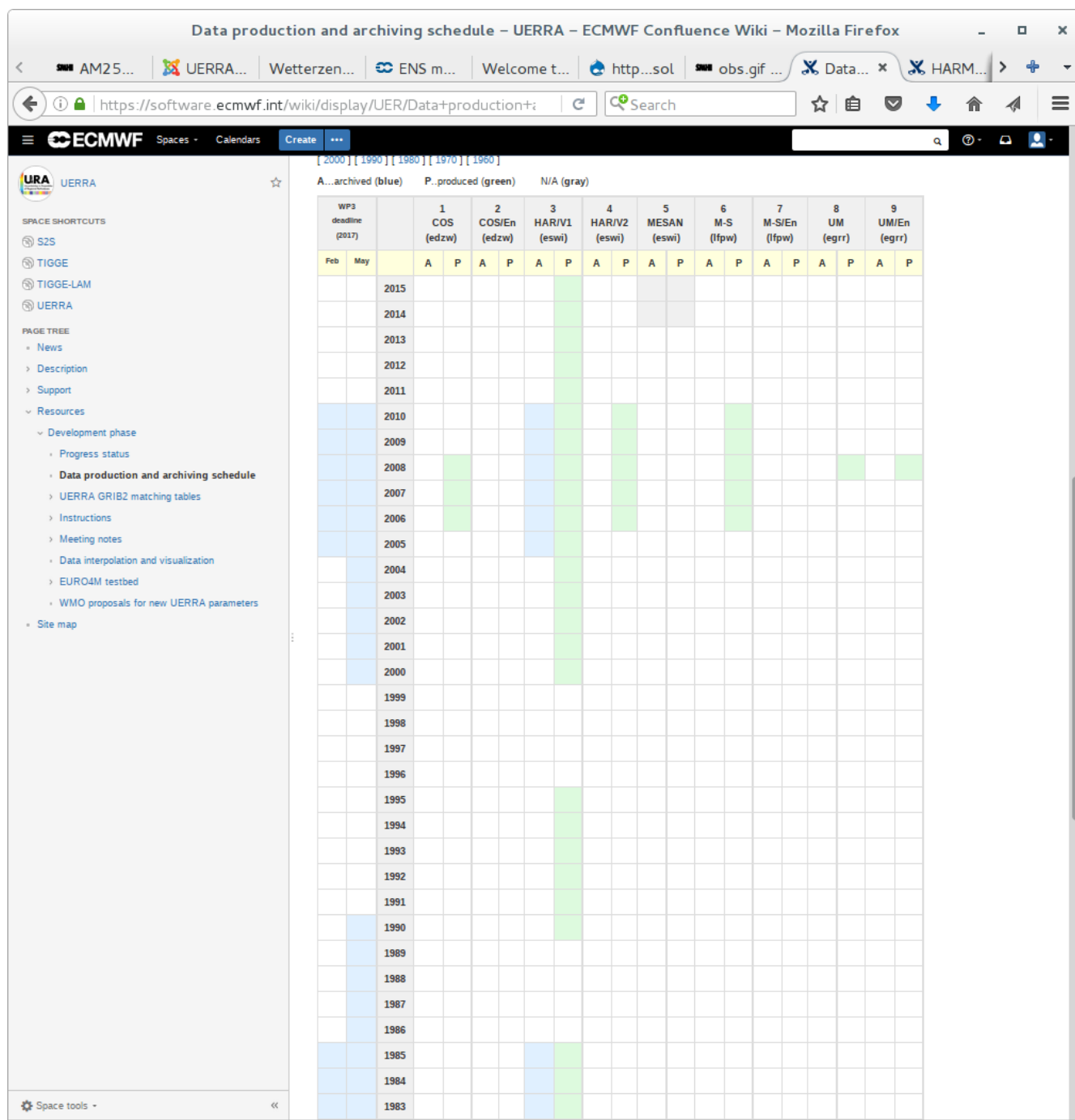*Figure 4. Status by year for the different RRA:s, produced and archived (blue).*
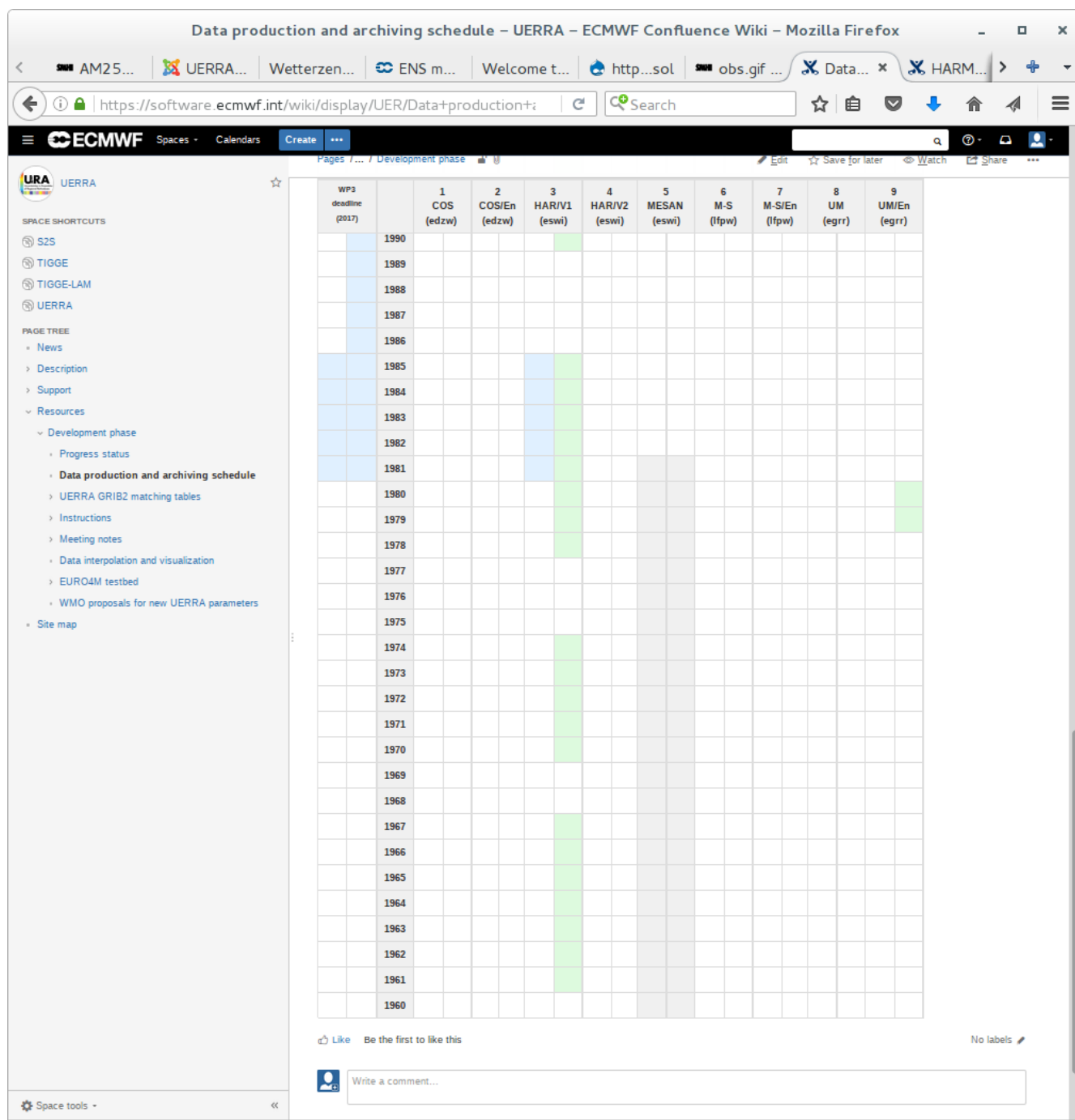
WP3 deadline (2017)		1 COS (edzw)	2 COS/En (edzw)	3 HAR/V1 (eswi)	4 HAR/V2 (eswi)	5 MESAN (eswi)	6 M-S (lfpw)	7 M-S/En (lfpw)	8 UM (egrr)	9 UM/En (egrr)
	1990									
	1989									
	1988									
	1987									
	1986									
	1985									
	1984									
	1983									
	1982									
	1981									
	1980									
	1979									
	1978									
	1977									
	1976									
	1975									
	1974									
	1973									
	1972									
	1971									
	1970									
	1969									
	1968									
	1967									
	1966									
	1965									
	1964									
	1963									
	1962									
	1961									
	1960									

Figure 5. As Fig. 4 but for early years.

UERRA General Assembly and Showcase event

The 4th UERRA General Assembly was held at ECMWF 21-23 November 2016 together with a Showcase Event 23-24 November. Many of the results and reports in this Newsletter emanate from these meetings.

Figure 6. UERRA General Assembly participants.

Figure 7. Participants in the UERRA Showcase Event, afternoon 23 – morning 24 November

Observation data rescue and data development

WP1 DARE efforts have been carried out both through scanned original data and from retrieving already digitised data. Some 9 M data have been digitised from some 15 countries in Europe and the Mediterranean area and most from after 1950. See the figures below for the counties where the data come from and then the numbers for each variable.
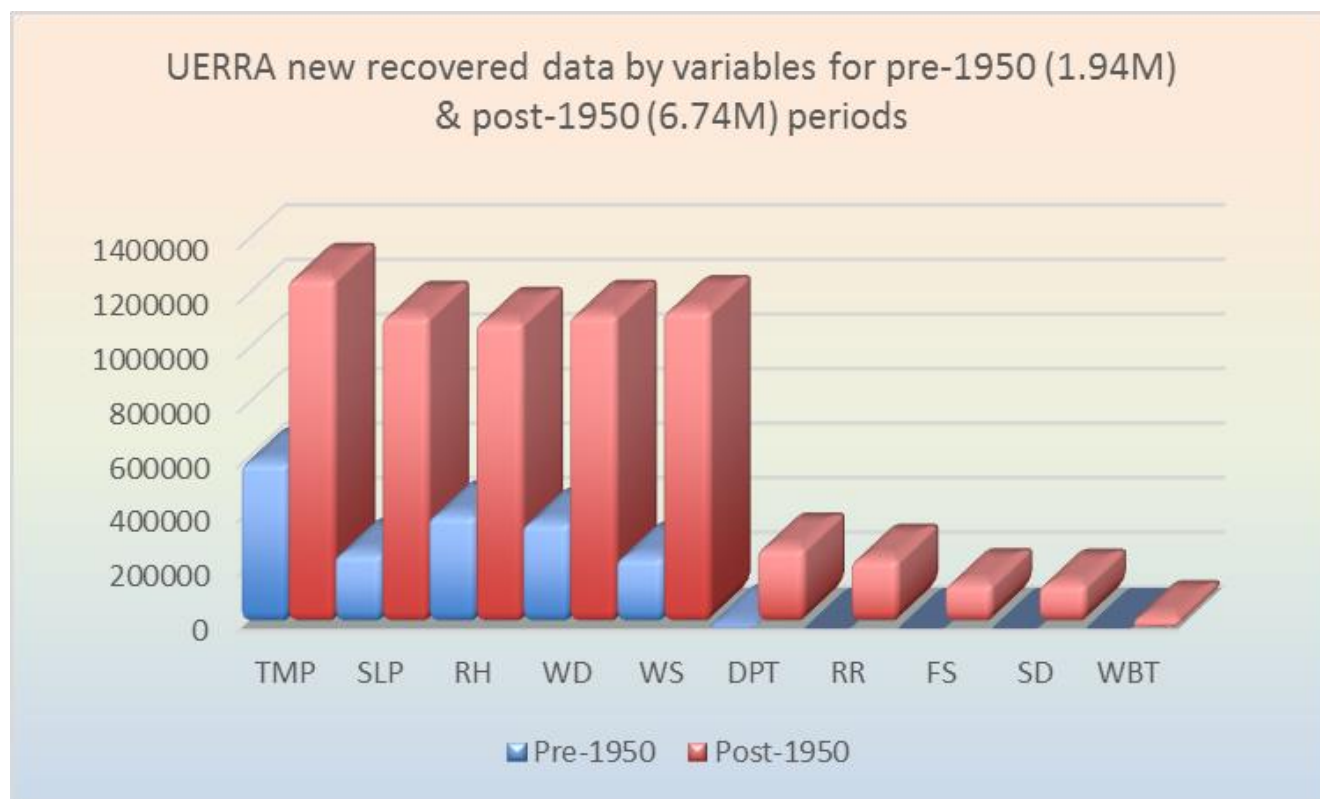

Figure 8. UERRA rescued data by variable (pre and post-1950)

In the current archives like MARS at ECMWF there are significant data holes in several countries in particularly the western half of Europe. Apart from the UERRA digitisation, a large volume of already digitised data have been received or retrieved from Sweden, Norway and Catalonia (see the three figures below).
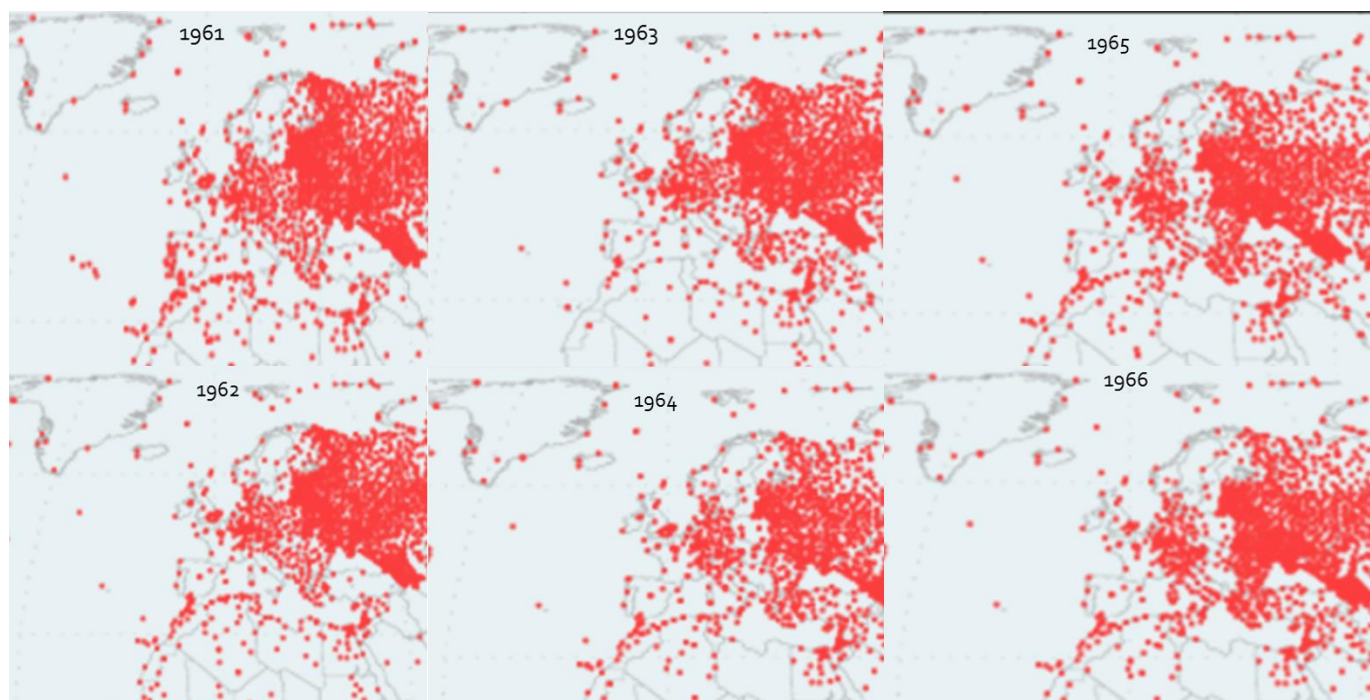

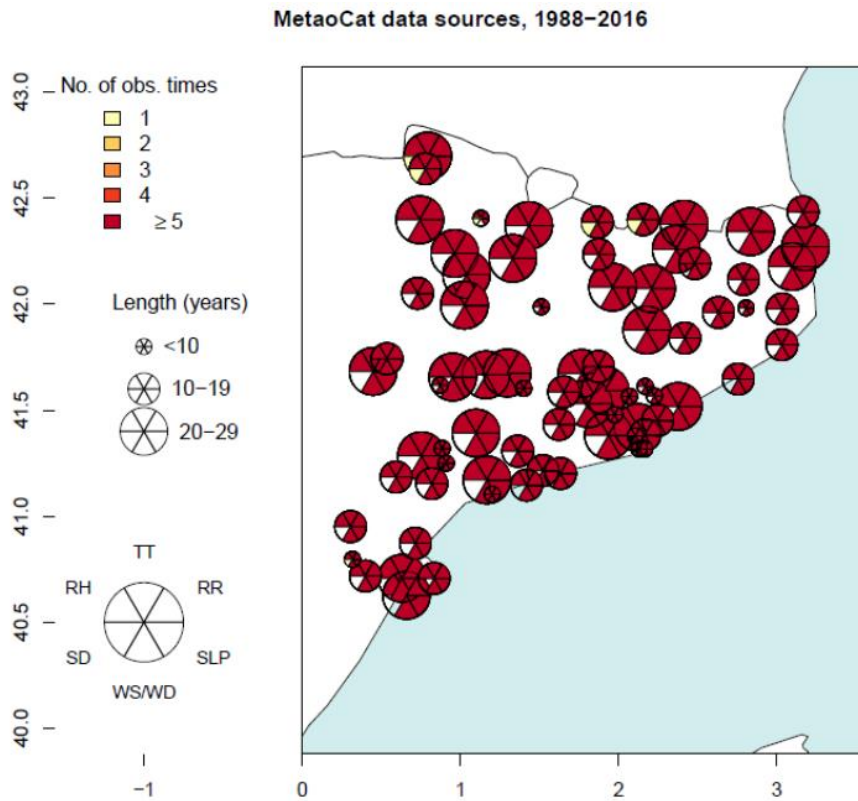*Figure 9. Key spatial data gaps in the early decade of regional reanalyses*

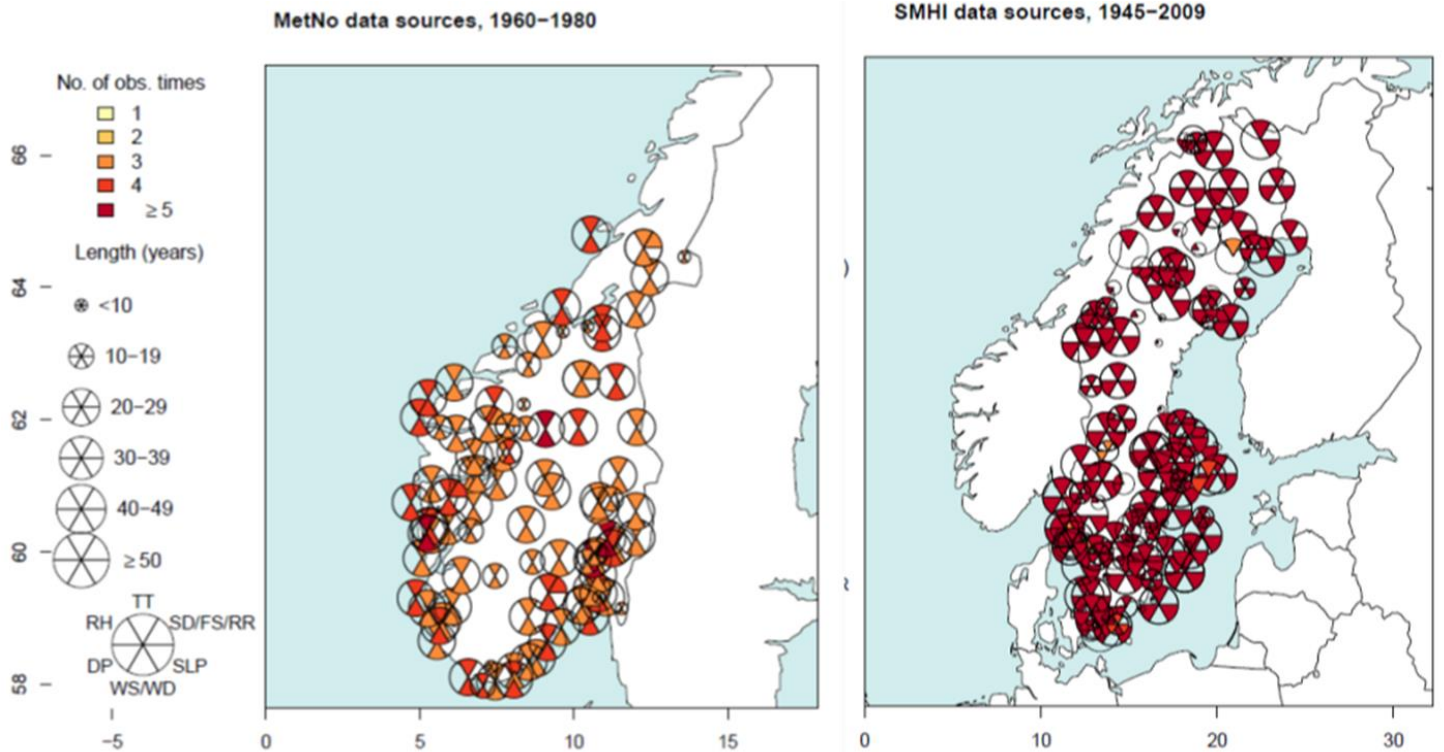*Figure 10. Received 129M data from Catalonia*

Figure 11. Retrieved 42M of data from Sweden and 7.2M from Norway

In total 178 M data have been processed with the vast amount of already digitised observations. An extensive quality control has been done on the digitised data, Visual Quality Control and Automated QC. A number of conditions are detected, like frequent occurrences of the same value and then spatial homogeneity is checked as well. In UERRA the novel aspects were related to also dealing with the sub-daily time scale. Visual checking is applied on values that are still suspect. A common problem is skewed values when the dates have slipped in the process. Almost all errors are detected and some can be corrected. Then there is the Spatial consistency check that can handle even more cases.
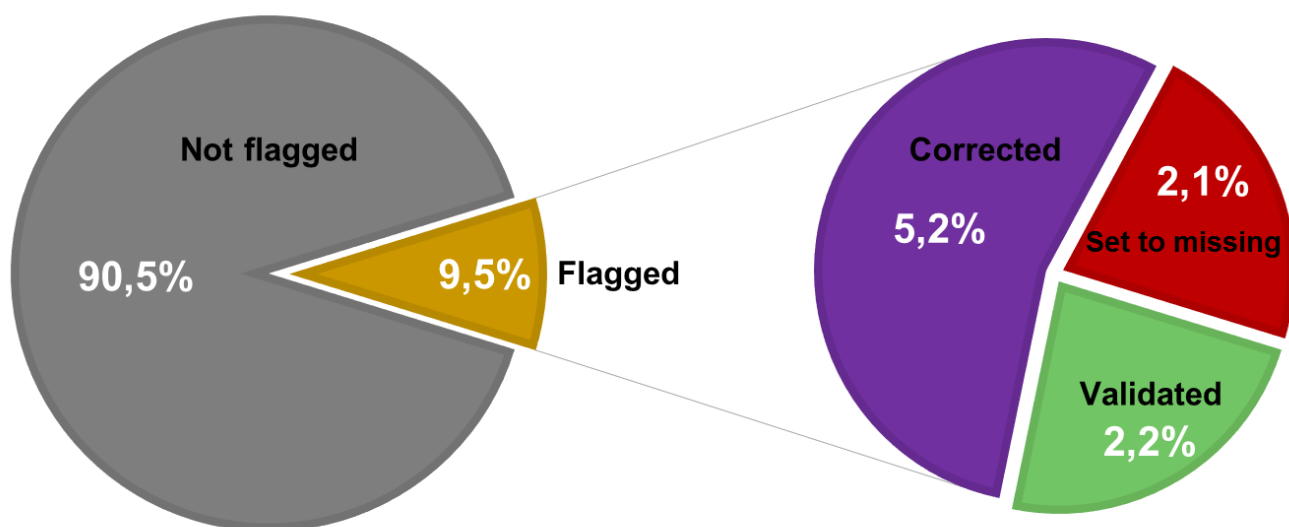

Figure 12. Rate of flagged data in the automatic QC and corrections.

On a slightly different strand, the hail climatology in Romania based on 105 stations was shown at the General Assembly. It is very much tied to high orography and for cases, on CAPE. It has been published in MWR (Burcea et al., November 2016 issue of MWR).

Developments of gridded observation data sets and uncertainties

E-OBS has evolved and particularly much more stations have been added from several countries and this risks to give a more inhomogeneous product. The interpolation methods have been developed with a gamma transform for precipitation. There is also a probabilistic product with 100 realisations that was shown in Deliverable D1.11. Apart from gauging uncertainties (see example below), the Ensemble mean can lead to improvements compared with the current E-OBS (see the figure below).The underlying grid has been changed and to a high resolution to avoid boundaries in user products. More parameters and high resolution orography are being brought in. The regressions will be done for more parameters. The dissemination of 100 members can be a problem technically and conceptually.

Another important factor for recent E-OBS versions is that the station density has increased dramatically for some countries when the complete stations network has been provided by the NMSs.

There are monthly updates to ECA&D station data and E-OBS but in addition there are also daily updates (not fully released yet) which gives a "first guess". It is valuable for rapid assessments of recent events.
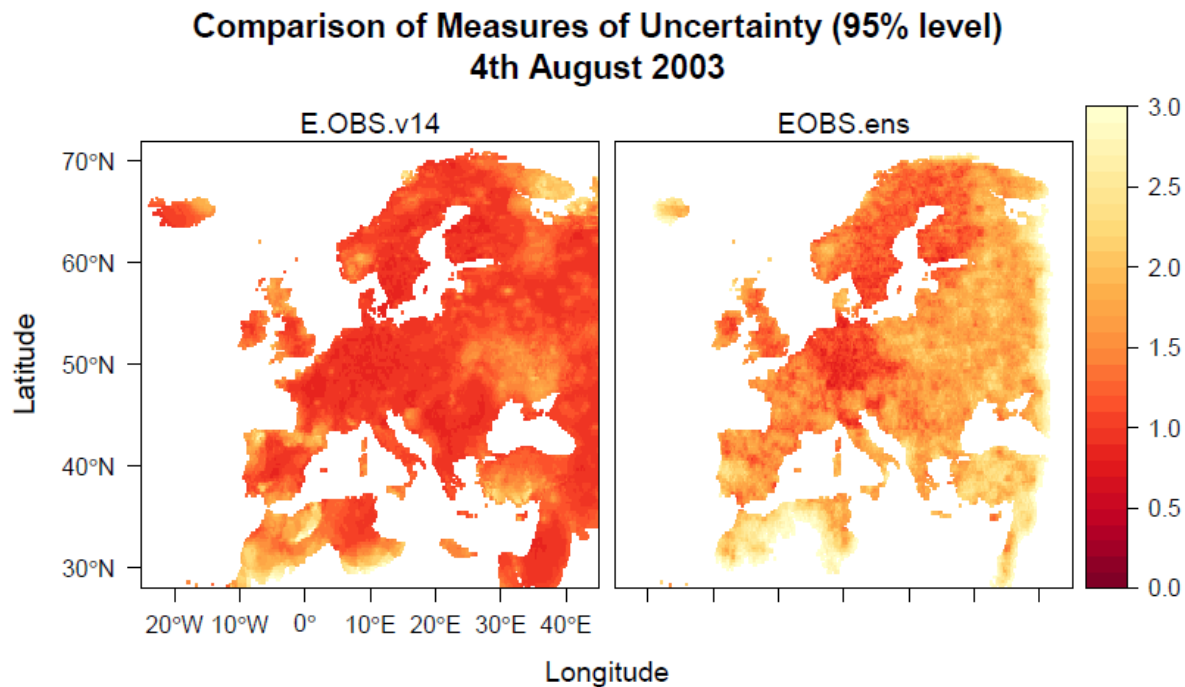

Figure 13. Uncertainty estimations of temperature from E-OBS ensemble for operational E-OBS (left) and new ensemble mean (right).

There is a stochastic precipitation multi-year data set for the Alps based on observation density and representativity and for catchments of varying sizes and it seems to be very realistic in its ensemble spread.
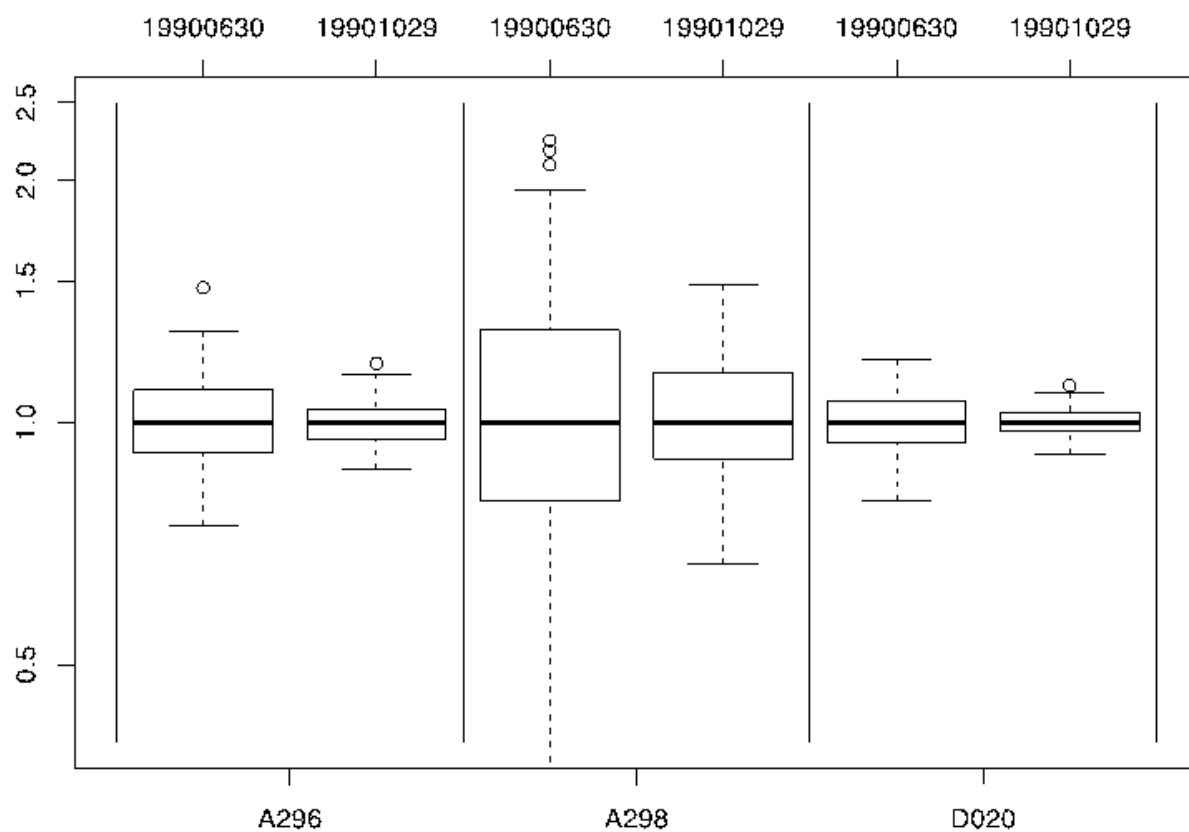

Figure 14. Examples of spread of catchment mean relative to median

Regional pre-operational Copernicus reanalyses

There are three different fully 3-dimensional reanalyses produced in UERRA and a fourth one analysing near surface parameters at high resolution, 5 km, instead of about 11 for the three upper air reanalyses.

Met Office Ensemble 4D-VAR.

The MO has developed and started to run the regional ensemble 4D-VAR for UERRA. The final configuration is 20 ensemble members at 36 km grid spacing and each member running its own 4D-VAR assimilation. Thereby the members can evolve differently due to non-linear growth of differences during the model runs. In addition perturbations of observations, sea surface temperatures and lateral boundaries are introduced. They are using ERA-Interim and not ERA5. This was shown not to have very much impact. The Ensemble assimilation has started from 1979 and a few years on including 1981. In addition the year of 2008 has been run in a preliminary version (with respect to observation processing and bias correction) and continuing for more years.

The deterministic high resolution 4D-VAR is running at 12 km grid resolution. It uses statistics from the ensembles of 4D-VAR at the lower 36 km resolution to enhance the background error covariances through the so called hybrid method.

So the ensemble spread is used both to estimate uncertainties due to paucity of obesrbations or errors in the observations, model errors and boundary conditions and also to use some of this information to provide situation (flow) dependent estimate of background errors. Examples of the spread are shown below for 2m temperature and 10 m winds.
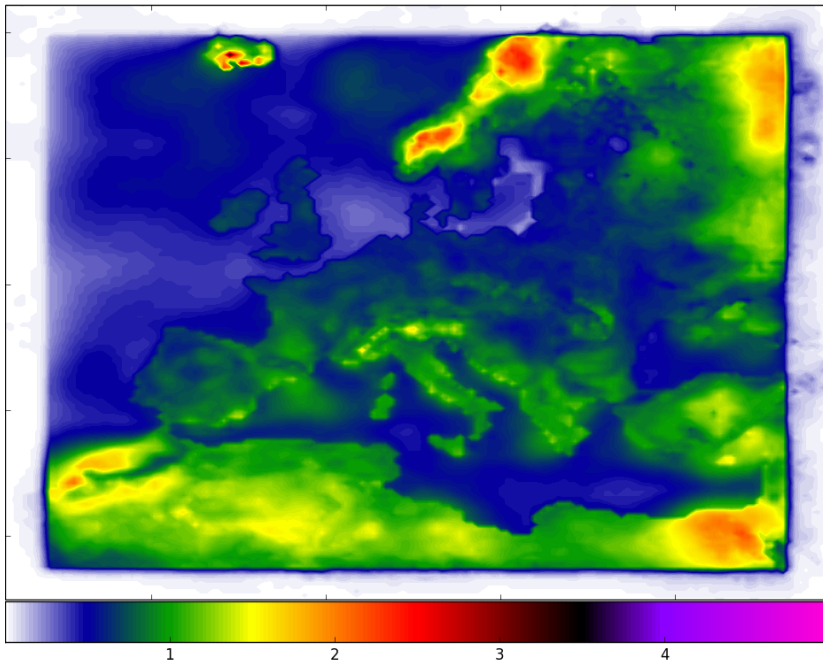
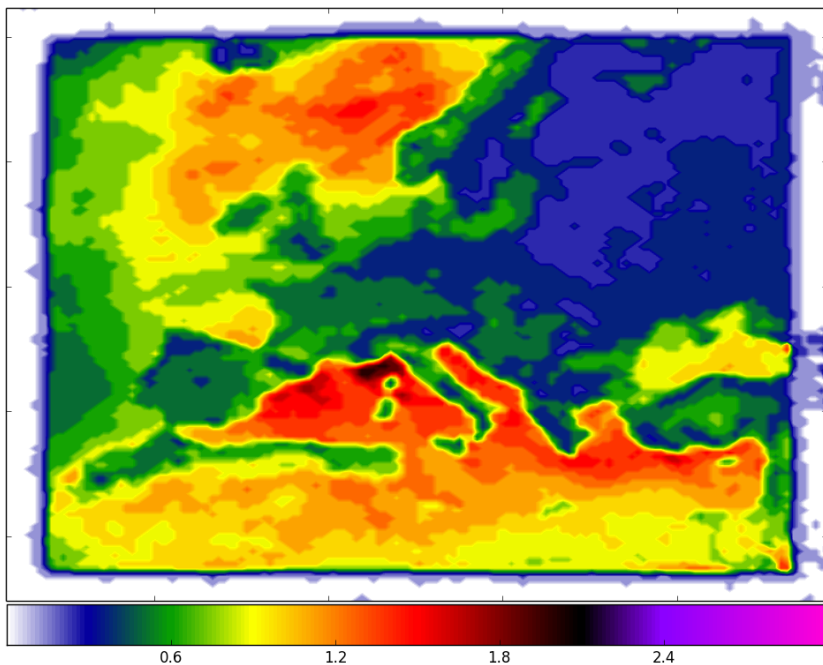*Figure 15. Spread in 2m temperature in the MO ensemble for March 1979.*

Figure 16. Spread in 10m wind speed in MO ensemble for March 1979

The suite encompasses 450 tasks per cycle so it is quite complex. The archiving and conversion to GRIB2 has been solved so the archiving in tte common MARS archive at ECMWF is also being done. Observations include the conventional and Satellite radiances and reprocessed satwinds and scatterometer winds and GPS RO and ground based. The ensemble spread and error and frequencies seem to be good.

A lot of work has gone into quality control of SYNOPs and rejecting suspicious ones. Outliers and suspicious ones are rejected. There is also a variational bias correction of satellite radiances. There will be a number of data denial experiments (particularly the Satellite one is important and demanded in the Description of Work in UERRA.

The ensemble and assimilation performance has been monitored against observations and for spread and shown to perform well. Below, the ensemble and mean errors against SYNOP 2m temperatures are show, together with the spread between members. It is still, in common with all other systems, in this measure under-dispersive. A recent **report** for UERRA is available now.
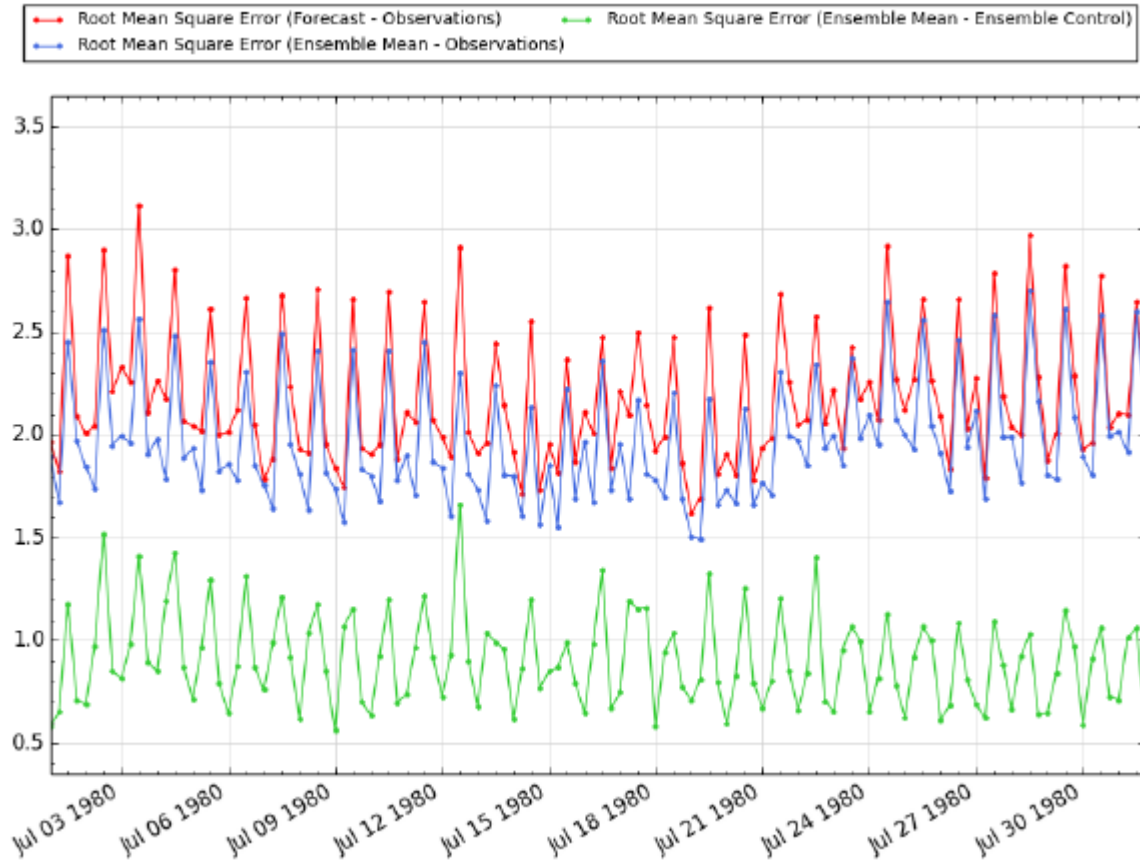

Figure 17. RMS difference of random member with observations (red), mean with observations (blue) and mean with random member (green), for temperature at 1.5 m in the model and observations normally at 2m.

The **SMHI ALADIN HARMONIE** has 5 years of reanalyses with the two model physics versions, ALADIN and ALARO. Then the years 2000-2015 with ALADIN is ready and 2005-2010 also archived in the UERRA MARS archive at ECMWF. 1961-1999 have between 6 and 7 years done per decade. Archiving in MARS of prioritized periods (for evaluation in UERRA) of 2005-10 and 1981-85 have also been completed.

Observation monitoring and verification of forecasts showed good performance and particularly temperatures are better (closer to observations) than ERA-Interim, see below. It is worth noting that the number of observations has increased, particularly for AIREP which did not exist in the beginning and the SYNOP stations were less and particularly some countries missing in the archives in the early 60's.
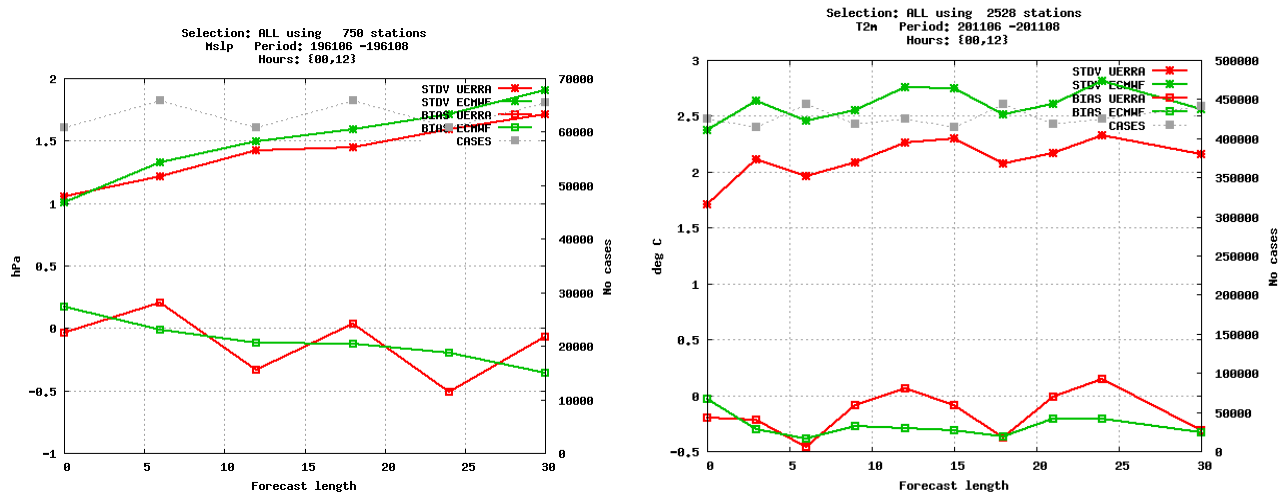
Figure 18. a) Standard deviation and bias between UERRA and SYNOP pressures and for ERA-Interim July-August 1961. b) for 2m temperatures.
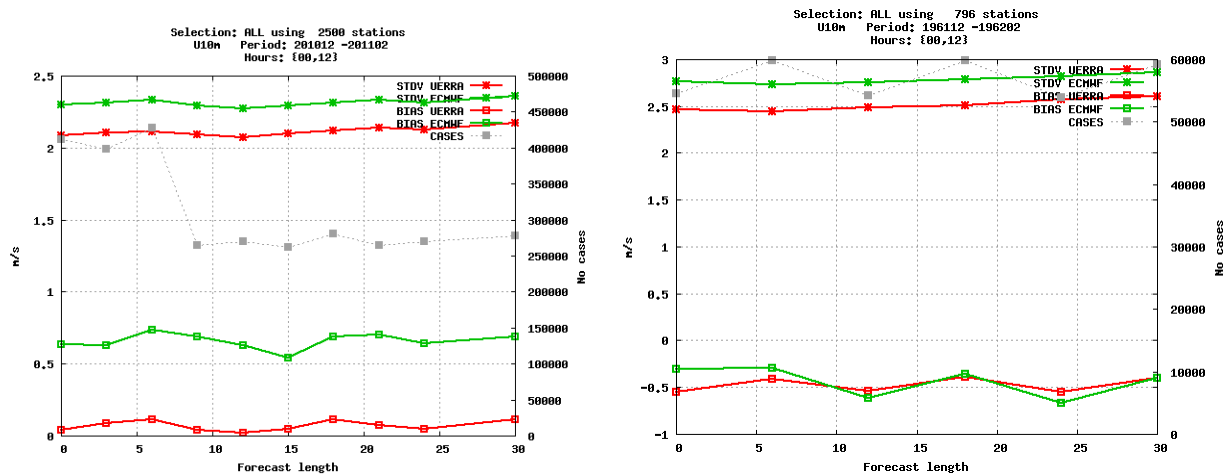
Figure 19 As for Figure 18 but 10m winds DJF 2011 (a). b) for 10m winds and DJF 1961

There are many different level types archived and the height levels are new and interesting for the wind energy:

- ➢ Soil: 2 parameters, 3 levels (only 6 hour forecasts).
- ➢ Surface: 29 parameters.
- ➢ Model levels: 4 parameters, 65 levels (only analysis).
- ➢ Height levels: 7 parameters, 11 levels.
- ➢ Pressure levels: 8 parameters, 24 levels.

There are analyses every 6 hours and forecasts up to 30 hours at 00 and 12 UTC and 6 hours length from 06 and 18 UTC. Every forecast hour is stored up to 6 hours and then every 3rd.

As an example of an application of interest for energy potential, the average winds at 150 above ground level from the SMHI ALADIN-HARMONIE reanalysis are shown below:
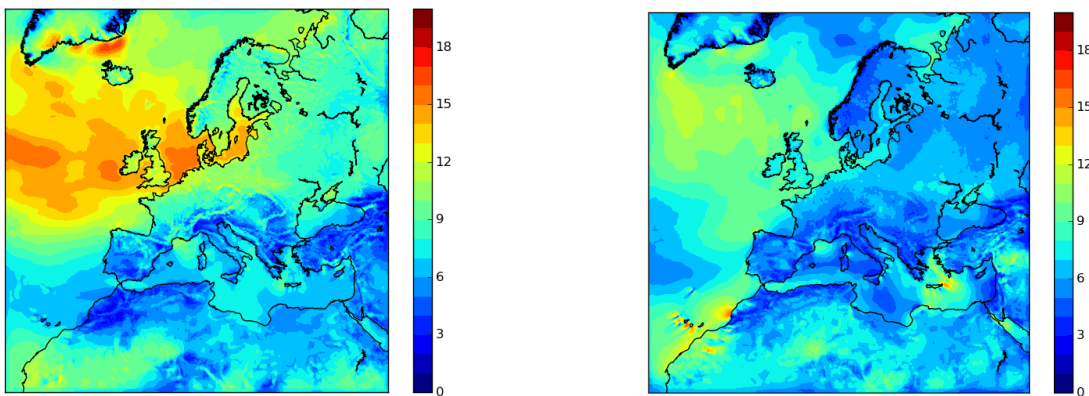

Figure 20. a) *Mean wind speed at 150 m January 2008. b) for July 2008*

The **Météo-France downscaling reanalysis** for 2006-2010 with ensembles has been done for both precipitation (see below) and temperature and humidity. The ranking of the members according to the observation distribution is shown below and is not perfect but reasonable but for the lowest and highest class. They are similar for 2006 and 2010.
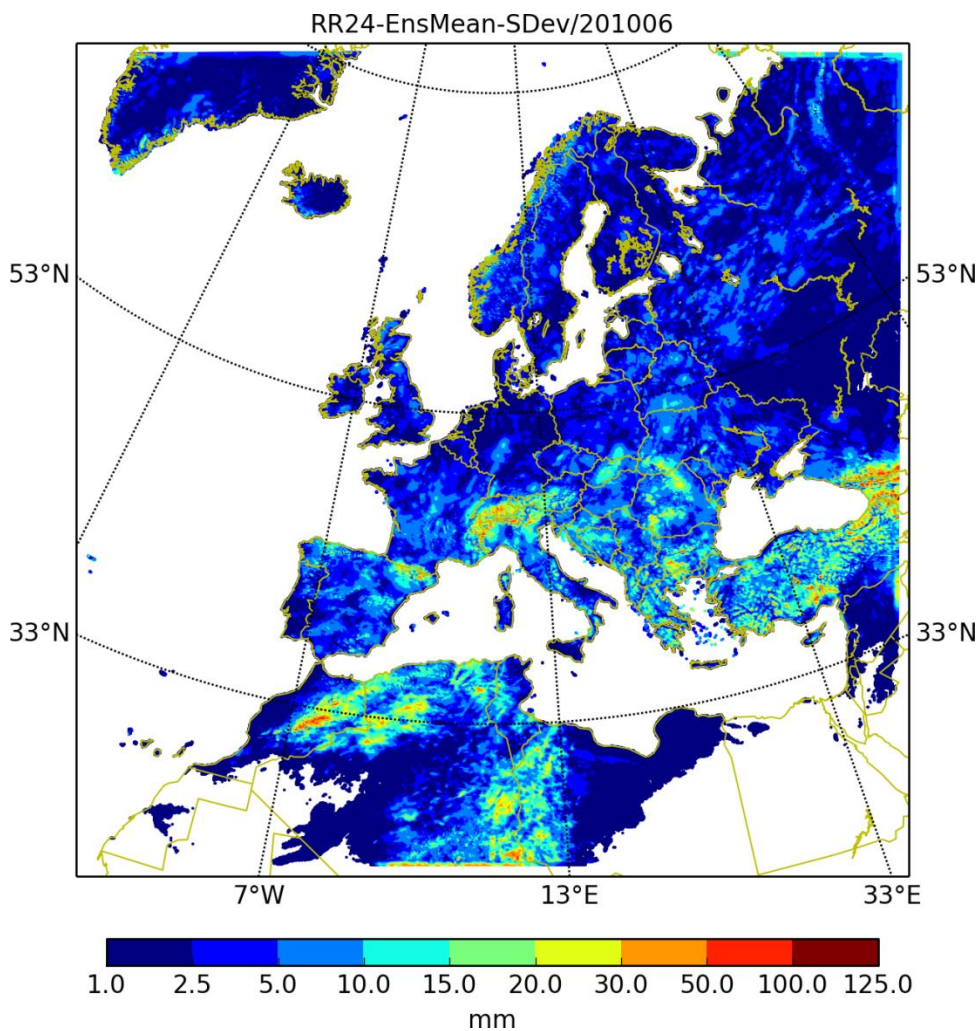

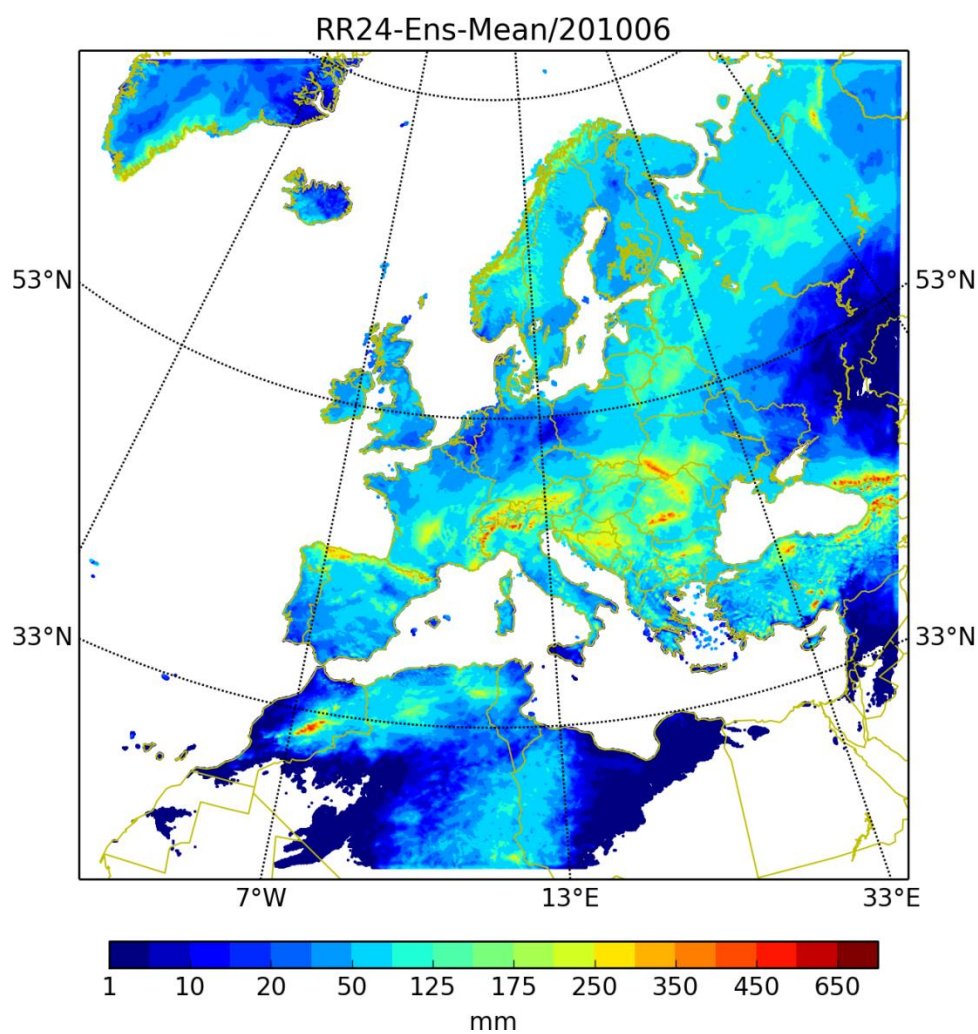*Figure 21. Ensemble spread for 24 hour precipitation for June 2010. Standard deviation.*

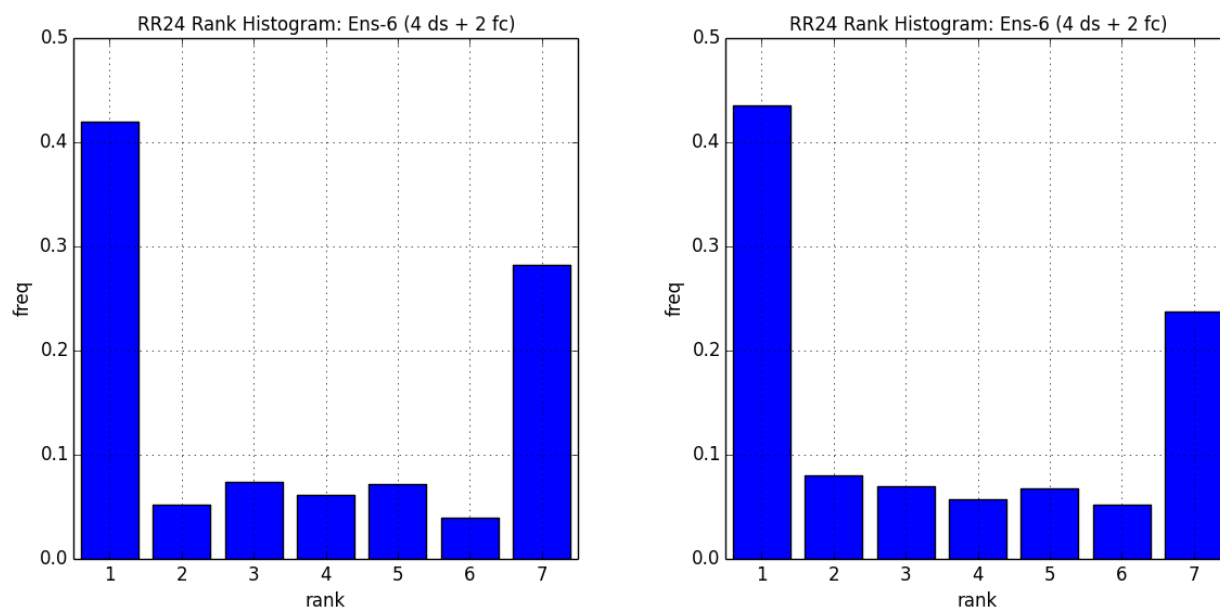*Figure 22. Ensemble mean for 24h precipitation for June 2010*

Figure 23. Rank histograms for January of 2006 with 6 members and b) January of 2010.

The number of surface observations has increased a lot for the 90:s and it may give a misleading signal. It has also been tested with low and high density networks. The downscaling production has started for the other periods.

The **SMHI** total cloud cover reanalysis has been and run for 2004-2008 on 5.5 km grid for the full European-Atlantic area as in EURO4M. (larger than UERRA). It is using Optimum interpolation but cost function based, to merge EURO4M background cloud fraction reanalysis and super observations based on weighted averages of data from CM SAF. (CLAAS-1 and CLARA A1). The difficulties have been in merging the different data sets and coping with stretches without observations (or processed observation rather). There are still some issues that are worked on and can be improved but the data set verifies well against independent data.

It will continue depending on usability of other data sets. One consistent new dataset that now is available for the years 1991-2015 and this period will be run with the new cloud fraction satellite product.
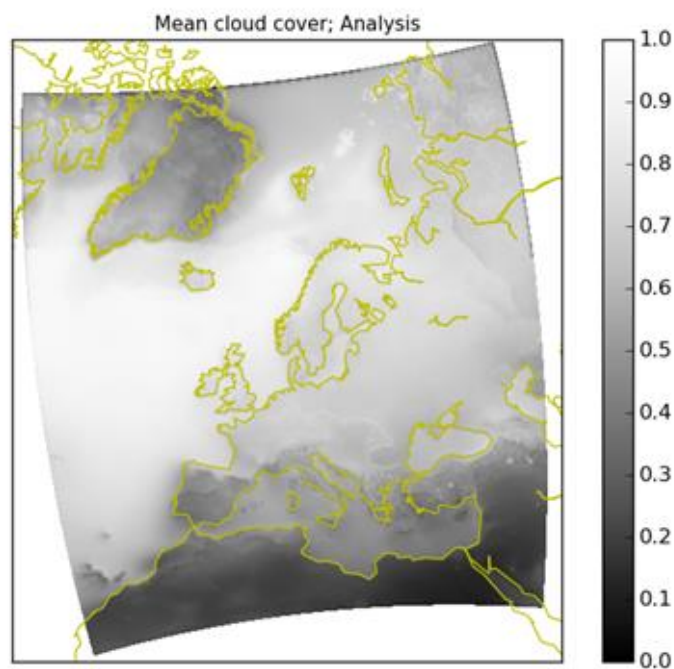

Figure 24. Mean UERRA Cloud cover analysis 2004-2009.

The **COSMO University of Bonn** (Hans Ertel Centre for Weather Research) ensemble nudging reanalysis has been run for almost all 2008 but intermittently stopped to sort out details of the model outputs to be archived. The five-year period will be resumed and produced and archived in the UERRA project.

It is a 20 member ensemble where each member uses nudging data assimilation with perturbed observations. All including the control are at 12 km grid resolution for the CORDEX Europe (4) domain. Conventional observations are used so no satellite data.

A range of probabilistic scores showed good performance and a paper has been accepted (Bach et al., Tellus 2016). Some examples are shown below.
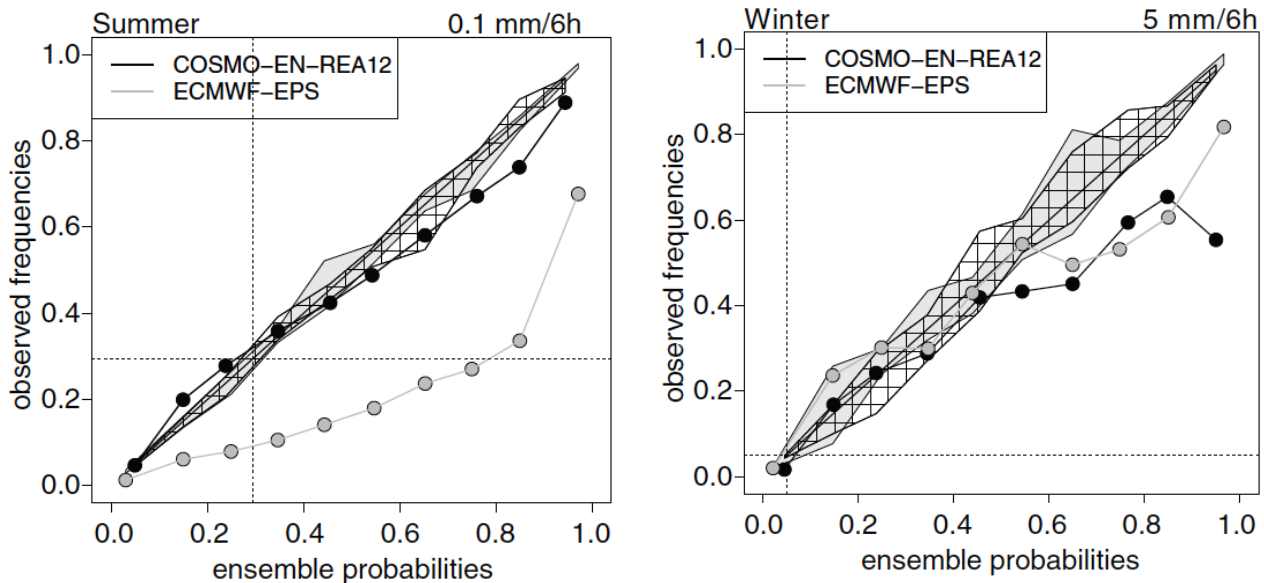

Figure 17. a) Reliability diagram of COSMO ensemble compared with ECMWF operational EPS for June 2011 and exceeding 0.1 mm/6h. The areas show consistency intervals due to limited sampling within which the system should lie (see Bach et. al for details). b) As for Fig. but December 2011 and above 5 mm/6 h.

The ensemble spread is expected to be a measure of the analysis and forecast uncertainty but usually it is lower than the forecast errors (in the case of forecasts from the ensemble). Examples of the spread and how the evolve over forecast length is shown below. Except for relative humidity there is a regular increase in spread with forecast time.
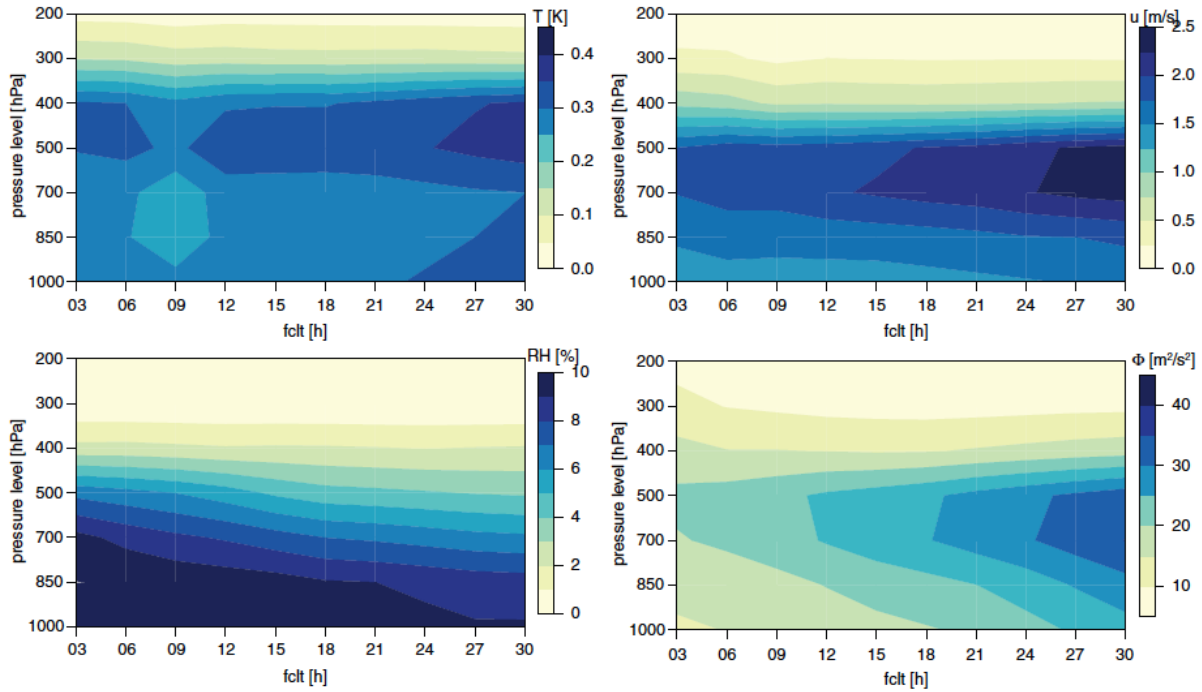
Figure 25. Ensemble spread (standard deviations) with forecast time for temperature, u component of wind, relative humidity and geopotential and pressure levels.

A **GANTT diagram** can be found on the UERRA web site with time line will be updated. At ECMWF there is also information on the wiki about UERRA, archive content and status of production. and the wiki at ECMW will have updated.

Uncertainty and quality evaluation and validation

The common UERRA evaluation package has been built at DWD and MET Norway which was developed from prior existing parts in R packages like HARP from the HIRLAM. A github repository is used where partners can contribute and the software can be shared and it is openly available.
It is used and tested on the samples of reanalyses (UERRA and prior existing ones) and with the data sets defined to be used in UERRA.

Since the UERRA production of reanalyses is well into production but there remains still quite a bit more to be done for some of them. Moreover the archiving in MARS lags behind the production so there have not yet complete data sets to compare the different UERRA reanalyses. There are some years, particularly 2008 where we have data but more will be available at the end of February. The tests of the evaluation software and comparison of some reanalysis data have been done on ERA-Interim, already available DWD COSMO RA (6 km CORDEX) and some SMHI ALADIN HARMONIE.

The different reanalyses have been intercompared and compared against observations, both conventional and then independent ones (not used in any reanlysis) from mast data.

Recently satellite data have been included in the evaluation and in terms of radiances which are measured and computed from the models in the reanalyses. They have been compared with satellite derived radiances, from CM SAF SIS data (Trentman et al). Even though the spatial patterns differed in the COMSO 6 km reanalysis compared with the ALADIN HARMONIE RA, the distribution of radiances were realistic in HARMONIE for high values and too little at low values (below, see Figure 26). For more information, see UERRA Deliverable 3.5 or Borsche et al. (2016).
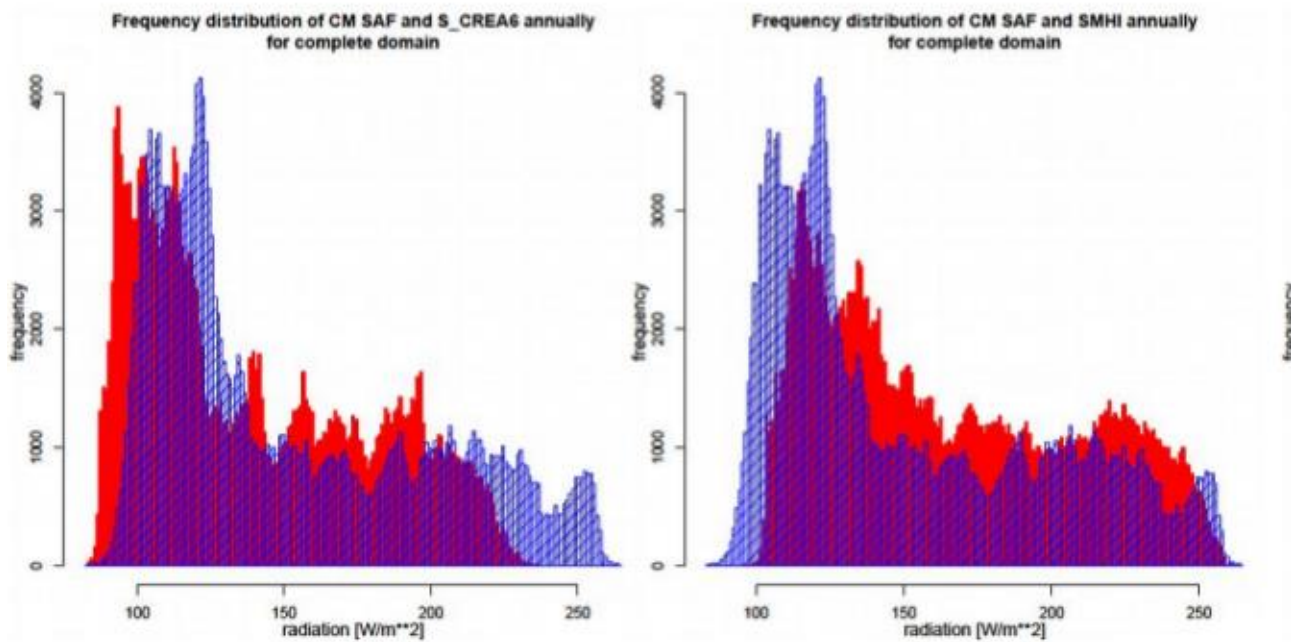
Figure 26. Comparison of the distribution of radiances in the COSMO RA and the ALADIN HARMONIE one in UERRA for the year 2008.

The Alpine Gridded Data set (ADG, Frei and Schär, 1998, Isotta et al. 2014) has been developed in the FP7 EURO4M project and further extended to provide probabilistic estimates. There are uncertainties in the observations, inhomogeneous distribution of them and in particular many uncertainties in the interpolation method parameters. The method model is assigned a distribution of parameters in an optimal way.

The distribution of rain accumulation in catchments is then possible to provide and such distributions are very useful when verifying deterministic reanalyses, to see if they are in the estimated range or not. An example is shown below in Figure 27.

The boxplots show the 25 and 75% percentiles and the ends of the whiskers the tails and dots show outliers.
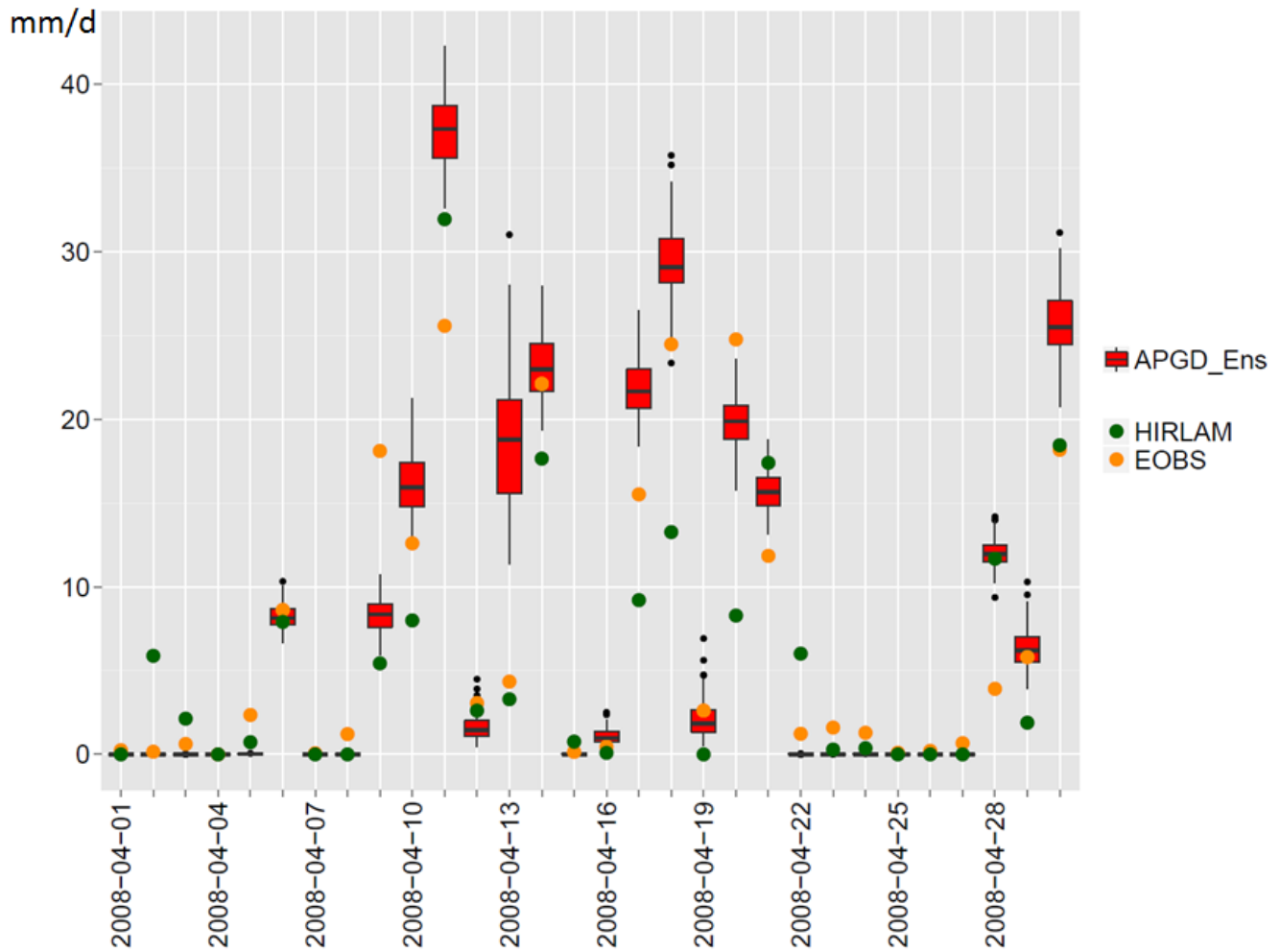
Figure 27. Daily precipitation in April 2008 in a part of the Ticino catchment area. The boxplots correspond to the probabilistic observational dataset, the green dots to the HIRLAM model (EURO4M) and the orange dots to E-OBS.

References:

Borsche, M., Kaiser-Weiss, A.K., and Kaspar, F.: Wind speed variability between 10 and 116 m height from the regional reanalysis COSMO-REA6 compared to wind mast measurements over Northern Germany and the Netherlands. Adv. Sci. Res., 13, 151–161, doi: 10.5194/asr-13-151-2016, 2016.

Frei, C., and Schär, C.: A precipitation climatology of the Alps from high-resolution rain-gauge observations. International Journal of climatology, 18(8), 873-900, 1998.

Isotta, F. A., Frei, C., Weilguni, V., Perčec Tadić, M., Lassègues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S. M., Munari, M., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Marigo, G. and Vertačnik, G.: The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data. Int. J. Climatol., 34(5), 1657-1675, doi: 10.1002/joc.3794, 2014