Seventh Framework Programme

Theme 6 [SPACE]

Project: 607193 UERRA

Full project title:

Uncertainties in **E**nsembles of **R**egional **R**e-**A**nalyses

Deliverable D3.1

Workshop on the definition of a common evaluation procedure

WP no:	3
WP leader:	DWD
Lead beneficiary for deliverable :	DWD
Name of author/contributors:	Michael Borsche
Nature:	Other
Dissemination level:	PU
Deliverable month:	3
Submission date: September 18, 2014	Version nr: 1

Minutes of UERRA WP3 Workshop (D3.1) On the definition of a common evaluation procedure Offenbach 26/ 27 June 2014

1. Table of Contents

1. Table of Contents..2
2. Scope of this workshop...3
3. Background Material...3
4. Agenda...3
5. Participants (UERRA / USERS)..6
6. Discussion on evaluation methods...6
7. On which scores to use..11
8. Blacklisting..12
9. Other topics...12
10. Presentation of users..13
11. Discussion with the users..15
12. Technical implementation..16
13. Telecon with Manuel Fuentes and Richard Mladek...18
14. Reference list..20

2. Scope of this workshop

Following the UERRA DoW and what was said at the General Assembly (GA) in Exeter in March 2014, work package 3 (WP3) arranged the workshop “On the definition of a common evaluation procedure” in Offenbach on 26/27 June 2014 to be applied to the regional reanalyses produced by WP2.

The discussion of common procedures and methods for the evaluation of regional reanalyses focused mainly on two subject areas:

- methods, metrics, and reference datasets
- technical details about topography, re-gridding, and storage and access of data

Users with interest in regional reanalysis data were invited and their interests have been discussed in detail.

The aim of the workshop is to arrive at an agreement on common methods and procedures taking into account user interests.

3. Background Material

In the months preparing the workshop, the online working document `summary_draft_Wp3_workshop_June2014` had been collaboratively developed, on which the discussions in this workshop were based on.

4. Agenda

Below the agenda for the workshop is included.

Uncertainties in Ensembles of Regional ReAnalyses

SEVENTH FRAMEWORK PROGRAMME

Funded by the European Union

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

UERRA WP3 Workshop on the Definition of a Common Evaluation Procedure (D3.1)

26 and 27 June 2014

Offenbach, Germany

Wednesday 25 June
Get together in the evening with anyone who likes to join

Thursday 26 June				
08:45	Registration and Coffee in signposted conference area			
09:00	Opening and Welcome	Frank Kaspar		
09:10	UERRA WP3 Summary of internet discussion	Michael Borsche		
09:20	Discussion: methods, metrics, parameters, resolution	all UERRA participants		
10:15	Coffee Break			
Parallel sessions for project team and users				
	Session 1		Session 2 (open to everyone @ DWD)	
10:30	Continued discussion	UERRA participants	Regional Re-Analyses of Observations, Ensembles and Uncertainties in Climate Information	Per Undén
11:00	Continued discussion	UERRA participants	Q&A of users	Per Undén and users

	Introduction of users (10min talk and 5min discussion)	
11:15	Use cases for regional reanalyses information in the context of renewable energy applications	Renate Hagedorn (DWD)
11:30	Reanalysis for regional climate modelling verification	Jennifer Brauch (DWD)
11:45	Regional climate data for biomass	Rolf Lessing (DELPHI IMM

	modelling	GmbH)
12:00	Requested climate data for agro-meteorology	Wolfgang Janssen (DWD)
12:15	Requested climate data for hydrometeorology	Annegret Gratzki (DWD)
12:30	Lunch Break	
13:30	Discussion on parameters, spatio-temporal scales of variability, questions addressed	together with users
14:30	Farewell to users and Coffee break	
15:00	Discussion: methods, metrics, parameters, resolution to provide answers to user questions	all UERRA participants
16:00	Common evaluation procedures (D3.2)	all UERRA participants
18:00	Soccer event for everyone interested	
20:00	Dinner altogether	

Friday 27 June		
8:45	Registration and Coffee	
9:00	UERRA discussion: technical implementation, data format, projection, domain, re-gridding	all UERRA participants
10:45	Coffee Break	
11:15	UERRA WP3 conclusions	all UERRA participants
12:00	Telephone conference with Manuel Fuentes and Richard Mladek about technical data issues	all UERRA participants
13:00	Lunch, end of workshop	

5. Participants (UERRA / USERS)

To the workshop all participants of WP3 and interested persons from WP1 and WP2 were present. In addition users with interest in regional reanalysis data were invited and joint part of the workshop. Below is a list of everyone present at the workshop.

UERRA participants

Lilo Bach, University of Bonn, Germany
Eric Bazile, Météo France, France
Michael Borsche, Deutscher Wetterdienst (DWD), Germany
Christoph Frei, Meteo Swiss, Switzerland
Peter Jermey, UK Met Office, United Kingdom
Phil Jones, University of East Anglia, United Kingdom
Andrea Kaiser-Weiss, Deutscher Wetterdienst (DWD), Germany
Frank Kaspar, Deutscher Wetterdienst (DWD), Germany
Jan Keller, Deutscher Wetterdienst, Germany
Cristian Lussana, Norwegian Meteorological Institute, Norway
Cornel Soci, Météo France, France
Ole Einar Tveito, Norwegian Meteorological Institute, Norway
Per Undén, Swedish Meteorological and Hydrological Institute, Sweden
Gerard van der Schrier, Royal Netherland Meteorological Institute, The Netherlands

Invited users (present)

Annegret Gratzki, Deutscher Wetterdienst (DWD), Germany
Renate Hagedorn, Deutscher Wetterdienst (DWD), Germany
Wolfgang Janssen, Deutscher Wetterdienst (DWD), Germany

Invited users (consulted)

Jennifer Brauch, Deutscher Wetterdienst (DWD), Germany
Nico Frischbier, Forstliches Forschungs- und Kompetenzzentrum (FFK)
Rolf Lessing, DELPHI IMM GmbH, Germany

6. Discussion on evaluation methods

Method A (feedback statistics)

In the feedback files, the assimilated observations (o), the background (b) also called “free forecasts” or “re-forecasts”, and the analysis (a) are stored. The main advantage is that the model parameters are brought into observation space with the observation operator so the comparison can be performed in the best way.

Usually, feedback statistics are a standard output of the data assimilation system and are frequently used by the producers for quality control. Note that o-a cannot easily be interpreted as the analysis is dependent on the observations. However, o-b is a good measure to start with as a means of comparison with independent observations (which is not strictly independent in all circumstances). From o-b we can easily compare bias and RMSE (root mean square error) from different reanalysis products. In order to draw significant conclusions one needs to take into account the different scales of representativity between observations and reanalyses. These scores are good for some but not all parameters.

For this method, radiosonde and aircraft data can be used because both data sources are assimilated by each reanalysis system. Other commonly assimilated observations are either not publicly available, which is a prerequisite for storing the feedback files on the MARS archive at ECMWF.

There are differences in the handling of the observations between the different reanalysis centres which complicates the interpretation of the feedback statistics. Pre-processing of aircraft data (filtering, plane dependent bias control) is complex and differs between the systems. In addition, the thinning of aircraft data is performed differently so that different systems might assimilate different observations.

Also for radiosonde data, the pre-processing and thinning differs between the reanalysis systems but not to such an extent as with aircraft data. Except for Météo France, the standard pressure levels are assimilated.

Conclusions:

We agreed to focus at first on the radio soundings and use all available parameters, which are temperature, wind speed, and relative humidity. We also agreed not to restrict ourselves on which heights to use, though our main interest is the troposphere (especially near the ground).

We agreed to share observations (o), the background (b), and analysis (a) at the observation point in the feedback files. These include the pre-processed observations. We also agreed to share the topography.

The producers agreed to store the feedback files in ODB-format and share them within WP3. This needs to be implemented by all producers.

Method B (compare to point measurements)

Comparing the regional reanalyses against station data poses many difficulties mainly due to representativity issues. But still users are very much interested in this comparison because they often have their own point measurement time series. So we need to be able to answer their question how good our reanalysis products compare against individual point measurements.

Another issue is that a very large number of ground station observations are assimilated into the reanalyses so that hardly any high quality independent data for an estimate of uncertainty are left. Especially the UKMO and Météo France products assimilate almost all available ground station data including those within ECA&D and parameters such as 2m temperature, precipitation, and 10m wind speed. Therefore, we agreed to follow two approaches by using not assimilated (independent) data (B1) and assimilated (dependent) data (B2) for uncertainty estimation.

The biggest issue with this method concerns the method with which the reanalysis is transformed into observation space in order to compare. Users have their own way to get model data into their observation space. For the comparison as discussed here, either a kind of post-processing or model output statistic (MOS) needs to be considered, or only a selection of stations with a large spatial representation, e.g., in flat terrain needs to be used.

When performing uncertainty estimation following method B1 using independent observations, Gerard van der Schrier suggested to use the data from the Cabauw, The Netherlands observatory and offered to establish contact to the responsible persons. There should be other such stations which measure at tall masts in Europe. Andrea Kaiser-Weiss, Michael Borsche, and Frank Kaspar agreed to check whether these kind of observations exist and are available for Germany.

It was discussed whether cloud height from ceilometer stations would be an option because data of these stations is (freely?) available. However, this parameter (cloud bottom height) is of minor user interest and in addition requires post-processing to interpret from the regional reanalyses.

In addition, it was discussed whether the radar-based precipitation product OPERA (http://www.eumetnet.eu/opera) could be used as reference data because it is not assimilated. However, it was explained by Frank Kaspar that the OPERA product is not a homogenized product yet. The effort to homogenize this product is too large for the UERRA project and hence cannot be used for evaluation purposes.

There was a larger discussion on how to determine and use Tmin and Tmax for evaluation. Phil Jones explained that there are some pitfalls when determining these parameters, especially when comparing against special station data, e.g., SYNOP [van den Besselar, et al., 2011]. Phil agreed on writing up a guide on best practices, including aggregation to daily means. Further clarification from Eric Bazile and Cornel Soci Tmin and Tmax analysis will be calculated as a pseudo analysis based on the T2m increment.

Conclusions

The summary of the discussion to this method is to divide the reference data into two groups of independent (B1) and dependent (B2) data because especially station data is assimilated in huge abundance. Independent data is hard to get at and needs to be followed up. It was agreed that from independent mast stations mainly the parameter wind speed would be analysed. For all the other station observations, it was agreed to analyse the parameters Tmin, Tmax, or Tmin and Tmax pseudo analysis where applicable, and number of days with threshold exceedance for temperature and precipitation.

Method C (compare to gridded data)

Validation against gridded fields which are spatially interpolated station observations. Several data products exist which cover the European continent or a sub-region thereof. The products themselves and the station data they are produced of need to be independent of the reanalyses.

It has been agreed upon that comparing reanalysis fields of precipitation against gridded fields is our choice for this method, because the spatial aggregation remedies the high local differences which make point comparisons very difficult. An ensemble of grids would be even better.

Members of this project have produced or gained very much experience with certain data products, which include Ole Einar Tveito who has access to a gridded data set for the Nordic region; Christoph Frei who is the maintainer of the Alpine Precipitation Gridded Dataset (APGD); and Phil Jones who has agreed to check on the availability of gridded data products for the UK. Christoph Frei added that it is planned to enhance available gridded data products such that they provide a probabilistic value, i.e they are planned to be produced based on ensemble methods.

With this method it is essential to keep track of which data each reanalysis system assimilates. For instance, the APGD is produced of about 6000 stations and Météo France assimilates only a sub-set of exactly these stations. Stations of this sub-set are located in the surrounding of the Alps but not within the Alps. The UK MetOffice plans to assimilate 24hrs precipitation sums of the E-OBS gridded data product which are partly made up of the same station data as the APGD but not entirely.

It was noted that it is essential to keep track of the version of the used data product because of the rapid development of some of these products as, e.g., E-OBS.

The parameter snow cover/snow depth and its assimilation was discussed. MESCAN does not assimilate any snow information. With SURFEX, however, snow parameters among them snow water equivalent are derived from the model output. Snow cover is very local, thus serves as check for MetOffice. SMHI assimilate snow depth.

During the discussion, it became obvious that an evaluation of precipitation would be particularly interesting for catchment mean conditions. For this purpose output from regional reanalysis and from grid datasets would have to be upscaled. Catchment polygons for European river systems are available from the EEA (European Environment Agency). Depending on catchment size, the upscaling would alleviate much of the scale discrepancies. For many users in the hydrological community catchment mean estimates are of primary interest.

Two further points were made relating to this discussion including

(1) when comparing two gridded data sets (i.e. gridded data set and reanalysis output) always upscale to the coarser gird (as had been done in the EURO4M approach) for not loosing information;

(2) for comparison not only precipitation is a useful parameter but also Tmax, Tmin, or Tmin and Tmax pseudo analysis as applicable, and Tmean;

Conclusions

The discussion about comparing against gridded data products revealed that it is important to keep track of the underlying station observations in order to decide whether the reference data set is independent of the reanalysis. It was agreed that precipitation for the uncertainty estimation of the reanalyses was the most useful parameter to use. Ole Einar Tveito, Christoph Frei, and Phil Jones agreed to offer their expertise and perform uncertainty analyses with available gridded data products. Next to precipitation, Tmin and Tmax are parameters to analyse for which the same considerations apply as outlined in Method B.

Method D (compare to gridded satellite data products)

DWD (Michael Borsche) will compare against satellite data products from CM-SAF (mainly radiation based products) and ESA CCI, e.g., GLOBSNOW were discussed as promising candidates. Michael Borsche noted that top of the atmosphere (TOA) radiances should be a good satellite product to add.

Remapping of the CM-SAF products was done for EURO4M with the tool CDO using “conservative remapping”. There was a discussion on the best practices of remapping and Phil Jones agreed to write a guide about that topic and Christoph Frei offered to be available for advice.

Method E (ensemble based comparison)

We agreed to revisit ensemble based comparison in the second part of the project, as all methods A, B, C, and D can in principle be done with ensembles. It has been noted that the ensemble realisations of the reanalyses will be in lower resolution than the deterministic runs. First results are expected next year, and production should start a year later.

Additional remark: Uncertainty measures can be calculated from an ensemble of reanalyses. In this case observations are usually taken as reference. When estimating uncertainty of an ensemble based input (i.e. ensemble reanalyses) ensemble skill scores are used together with a deterministic reference which in our case are usually observations. It has been shown, that neglect of uncertainties in the reference can lead to overly pessimistic measures of skill in probabilistic evaluation. Proposals have been made how some of the available verification concepts can be extended to the case with uncertain observations. (Bowler 2006; Candille and Talagrand 2008.)

Method F (User related models)

The development of SURFEX at Météo France was acknowledged. Jan Keller has mentioned that the group TR32 in Jülich/Cologne might be interested in providing such a user model and contributing to the project.

Method G (re-forecast evaluation)

During the discussion “forecast” or “re-forecast” evaluation came up several times, however, we did not agree on whether to follow up on this method of uncertainty estimation. The reason is that different data assimilation systems have their own skill scores they look at which might not be comparable across systems. For more details on which skill scores to use refer to Section 7.

We did agree on though, which re-forecasts we would like to store. For further details on storage requirements and ideas refer to Section 13.

7. On which scores to use

The focus of this workshop was on which parameters, data products, and methods to use for uncertainty estimation of the regional reanalyses produced in UERRA. There was some discussion on which scores and skill scores to use. Generally, the bias and RMSE are good starting points but are not the optimal choice for all parameters of interest and might not capture the whole spectrum of uncertainty. There are many other scores which are designed to answer specific scientific questions and should be applied to certain parameters only. For a guide on which questions should lead us we can refer to the user questions in Section 11. Some of them were named during the discussion but not in an exhaustive manner. For different parameters commonly used skill scores include for wind gust the fractional skill score (FSS), and for precipitation the Stable Equitable Error in Probability Space (SEEPS), Equitable Threat Score (ETS), and Hansen Kuipers skill score. There are robust skill scores especially developed for more extreme events (consider whether symmetric distribution) which include the Symmetric Extremal Dependence Index (SEDI) for rare binary events, see [Ferro and Stephenson, 2011].

8. Blacklisting

Blacklisting refers to deliberately excluding observations from the data assimilation in order to have them available as reference for the verification. However, following this approach will probably decrease the quality of the reanalysis because the error introduced by the data assimilation is a function of the station density. Blacklisting of some stations might lead to results which have a significant random component. On the other hand, if blacklisting does not make a difference, then the stations in question were very close to each other contributing similar information content to the system. In those cases the data assimilation system itself would decide not to include both stations. In addition, using one of those stations as reference does not provide an independent observations.

In the discussion it was found a common blacklisting experiment would be desirable, covering a five year period. It should be run as an extra experiment, in which not all observations are used, but some are blacklisted and hence can be used for the evaluation method B1. This blacklisting experiment shall be checked against available resources and then agreed upon by the reanalysis producers (Per Undén, Eric Bazile, Peter Jermey, and Jan Keller).

An open question concerns where to store the commonly used and blacklisted observations? Two options were discussed, either archiving these observations at KNMI or ECMWF.

9. Other topics

- We need to capture which input observations have been used, the forcing history, trace gas concentrations, aerosols, vegetation, land cover, and determine whether they are constant over time, we use a climatology, or whether they are described as a function of time.
- An advice for reanalyses in general is to include an extra day, or as appropriate, for the start and end time of a reanalysis run, in order not to have missing fields. For instance, if the assimilation window starts, e.g., at 21 UTC but the starting point would be 01 January at 00 UTC, the first value would be missing. The same is true for forecasts: if a forecast of x hours is desired for 01 January 00 UTC then the reanalysis needs to start that x hours (plus the length of the assimilation window) before. In addition, the extra day is particularly relevant for the evaluation of precipitation, Tmin, and Tmax, where a daily total/extremes are usually not taken over the midnight to midnight period.

10. Presentation of users

Per Unden presented the UERRA project to a broader interested audience with an overview talk. In addition there were a number of users, persons from or heads of other sections of DWD responsible or interested in products from regional reanalyses. They were mainly representing climate and energy, hydrology, agro-meteorology, and forestry. These users introduced their work and interests in short presentations.

Renate Hagedorn: wind energy applications

Renewable energy is very important in Germany now and there are government programmes to provide information to companies. Several projects are followed at DWD in order to better estimate power production with threshold power curves. First, the wind energy potential is of interest, then in more developed areas optimized integration into the power grid, and predictability of production are of interest (it becomes more and more important to know when the power is available, for shutting down or switching on other energy productions).

Reanalysis data can help the transformation operators to develop a strategy in accordance to meteorological conditions. At the moment the estimation of wind speed distribution is performed relying on the Weibull distribution of measurements and forecasts. The main parameters of interest for a better estimation of wind speed are roughness of terrain and the distribution of wind direction. Reanalysis data was pointed out to be specifically used in wind power production

- for verification of datasets to improve forecast, need for “unobserved” variables like wind speed in 100m height;
- to optimize the power prediction function (due to sparseness of data, Weibull parameteres are currently used in training phase) – better wind fields at 10 m height wanted;
- for better estimation of wind speed and direction variability at 50m to 200m height (frequency of so-called ramps which are on sub-hourly scale, about 15 min), information on predictability and uncertainty (quality).

Rolf Lessing: biomass modelling

The presentation was provided by Rolf Lessing. As he could not give the talk himself, Andrea Kaiser-Weiss presented it. A biomass process oriented model was introduced that works on a very high spatial resolution in the order of 1 km^2 and daily means. Forcing data is mainly taken from satellite remote sensing products and climate station data records with which net primary production (NPP) is calculated. Since NPP is highly variable and sensitive to the input, regional reanalysis products might serve as a better forcing data product. Especially, highly resolved precipitation data is needed for the model. In the discussion it turned out, however, that the horizontal resolution requirement would call for a MOS (model output statistic), and that the desired cut-off values at one distinct point probably have similar uncertainty or error for station data, satellite products and the regional reanalysis.

Andrea Kaiser-Weiss agreed to keep in contact and inform Rolf Lessing about the EURO4M project, MESCAN by Météo France, datasets as a starting point for initialization, and progress within UERRA.

Jennifer Brauch: Regional Climate Modelling (RCM)

Personnal communication prior to the workshop with Jennifer Brauch clearly pointed to RCM needs of verification data sets on the CORDEX grid. Since the regional climate model is based on a version of the COSMO model, as is the COSMO reanalysis, it would be best to compare to reanalyses based on different models.

Nico Frischbier: Forestry

In an email prior to the workshop Nico Frischbier informed us on how he would decide whether and which reanalysis product he would use in forestry applications. In order of importance:

- trends and frequency distributions have to agree with station values
- high spatial resolution
- plausibility and consistency, high information content
- correlation of reanalysis with station series (i.e. rain at the correct day)

Wolfgang Jannsen: Agro-meteorology

Also in agro-meteorology modelling of biomass with a process oriented approach is carried out. Soil moisture is a very important input variable for their models. Other parameters include dew point, radiation, precipitation, and wind speed on an exact point-wise spatial resolution. Temporal resolution does not matter that much, but hourly would be perfect and six hours may be sufficient. So, for agro-meteorology the frequency and trends are the most important quantities (timing is not). However, time series at a specific point is not what regional reanalyses can resolve on their own but which could be solved applying a MOS (see biomass modelling above).

Annegret Gratzki: Hydrology

Hydrologists are interested in catchment areas down to the smallest scales. One challenge is that catchment areas cross country borders which makes it very difficult to create high resolution data products as is being done in the E-HYPE project. The interest in regional reanalyses is that they cover the complete European area with internal consistency and cover a long time period. The free availability is a big plus compared to some observations. If the uncertainty of regional reanalysis is in general equal or smaller than the used observation products then the reanalysis products is preferred. In UERRA there will be only a limited ensemble of hydrological realisations (WP4) through different RA input. The uncertainty then depends on catchment sizes.

11. Discussion with the users

For the renewable energy community, especially in wind power generation, high temporal resolution (about 10 min) and moderate spatial resolution of 10 km is of interest. Long-term change in wind speed (climatology) is of equal high interest. There are some well funded people in the community who have sophisticated knowledge to interpret feedback statistics, and the others would benefit from advice and guidance on the interpretation of uncertainties. The ensembles of the reanalysis would also be used and interpreted. If point information is needed either a MOS-system is used directly or other post-processing methods applied. For wind, point-specific forecasts from the 2.8 km COSMO-DE grid are produced.

User main interest: parameters

- temperature (Tmean, Tmin, Tmax);
- relative humidity (dew point);
- precipitation (and phase, which can be provided by regional reanalyses but is difficult to verify);
- soil moisture as a function of soil depth;
- wind speed and direction – model levels are okay, but discrete height levels (at 100 m above real topography) favoured;
- (Relative) sunshine duration, solar irradiance (direct and diffuse – total of all wavelengths),
- Clouds;
- snow cover, snow depth and/or snow water equivalent;

User main interest: uncertainites

- in principle user look for a reference data set which is simple to interpret;
- error estimates (for each parameter) as a function of spatial and temporal scale, and dependent on meteorological condition, e.g., convective and stratiform precipitation;
- suggestion: start with something reliable and then zoom into smaller scales, certainly valuable for different regions (flat, mountains, etc);
- ensembles as a measure to judge quality of data by providing some uncertainty characteristics (e.g., spread);

- for now: give most realistic member (many users want only one case); in a few years maybe interest extends to ensemble;
- hydrological users are starting to use forecast scores;

User main interest – other

- own comparison with own gridded fields planned;
- comparable successful applications to show to new users;
- value in simple examples demonstrating uncertainties, demonstrating added value of ensembles (ensemble derived uncertainties), or demonstrating applicability with extreme events (e.g., SURFEX);
- time series, climatology (frequency distributions) at grid points, trends in means, and trends in extreme values
- 'extremes' with respect to applications (which is a non-linear function of meteorological parameters)
- aiming for ca 1x1 km or below

12. Technical implementation

Technical details were discussed on Friday, 27 June 2014 including the following topics:

ranging from which fields to store in which format, to how to best perform re-gridding and where to do it (at ECMWF or locally), re-forecasts, and aggregation.

- After re-capping what users want and what data we have and what is possible we (WP3) have to consolidate: „Which questions we want to address (parameter-dependent, scale-dependent) " -> after that we look for the most useful quantification (score)

Aggregation to daily values

The discussion about aggregation focused on whether to provide an aggregate in addition to the reanalysis output and how to perform the aggregation. After intense discussions it was agreed to provide a daily aggregation of precipitation from the reanalysis re-forecasts. Phil Jones argued that it would be best to prepare a guide on best practices of aggregation and have those aggregated products available, but also let the user have the possibility to produce their own aggregation on MARS. It was agreed that we provide a daily mean of precipitation and temperature, Tmin, and Tmax between 6 UTC to 6 UTC of the next day. When comparing against observation data this is the most commonly used aggregation, however, there are different conventions, for instance in some countries between 7 UTC to 7 UTC, etc.

Given the 24 hour forecasts starting every 6h, there are different options to construct daily precipitation from reanalyses. Landelius tested forecast length.

Soil Moisture

The soil moisture outputs (and archiving) were discussed a lot. The exact depth varies between models and may be hard to compare. For instance, in the TIGGE-LAM project they are re-calculated to common layers.

Gerard van der Schrier pointed out a recent paper [Qui, at al., 2014] which evaluates the information content received from the very narrow penetration depth of microwave remote sensing, as for instance that from satellite, and concludes that it is possible under some circumstances to relate the superficial soil moisture observations to more relevant depth of soil moisture.

Analyses and Re-forecasts

The discussion about which levels, parameters, and time steps of the analyses and re-forecasts we and the users wish to be stored ended in the following conclusions.

Re-forecasts should be stored starting from 00 UTC and 12 UTC with time steps in hours of T+1, 2, 3, 4, 5, 6, 9, 12, 15, 18, 21, 24, 27, 30; and starting from 06 UTC and 18 UTC with time steps in hours of T+1, 2, 3, 4, 5, 6. This shall include the complete time period of the reanalyses as produced by each centre and include surface and standard pressure level fields.

The analyses are to be stored six hourly at 00 UTC, 06 UTC, 12 UTC, and 18 UTC for all standard pressure levels and all model levels in the troposphere. The COSMO re-analyses are to be stored hourly at all standard pressure levels.

Considering the parameters to store for both the analysis and re-forecast we had a look at what is stored by the UK MetOffice EURO4M reanalysis and decided that this list of parameters would be a good start for UERRA. The list includes seven parameters at model level, five parameters at pressure level, and 51 soil, surface, and near surface parameters.

Michael Borsche will prepare a table summarizing which parameters have been stored by the EURO4M reanalysis of UK MetOffice and which parameters are stored for ERA-Interim. The group needs to decide on which of these parameters we would like to store in the end and then figure out whether this can be done in grib2 format. If certain parameters are not included in the grib2 format definition, we need to contact WMO, probably via Manual Fuentes (see Section below for further information).

Comparison table

It was agreed that we will prepare a comparison table of all reanalyses produced in the UERRA project which shall be similar to that found on the website http://www.reanalysis.org for global reanalyses. Also the users were asked whether they would use this kind of table and we received positive feedback. Michael Borsche will be responsible to collect input from the project participants, include the user feedback, and distribute the final table to the UERRA web site and in a mature state to http://www.reanalysis.org.

13. Telecon with Manuel Fuentes and Richard Mladek

A telephone conference was arranged with Manuel Fuentes and Richard Mladek at ECMWF about the data to be stored as discussed by WP2 and WP4 and archived in WP4.

The format of choice at ECMWF to store model output is GRIB2. The producers of the reanalyses need to make sure that all parameters and their respective projections can be stored in GRIB2. If parameters are missing in the GRIB2 format definition we would like to archive, we need to get in contact with WMO, through Manuel Fuentes, to add them.

Furthermore, Manuel Fuentes and Richard Mladek agreed to look after the implementation of the Lambert Conformal projection into GRIB2 format, which is the projection of SURFEX. Additionally, the soil moisture field, from SURFEX but maybe also from other reanalysis, is a field with variable depth (number of layers) and it needs to be possible to archive just like that.

All observation based data products, including the feedback files are to be stored in ODB2 format which is machine independent. Since the MARS archive is in principle freely accessible, all observation based products need to consist of free data. This will restrict especially precipitation data products and our ability to analyse precipitation feedback statistics. Producers need to be careful to think about that when archiving feedback statistics.

We have iterated with Manuel and Richard the issue with daily and monthly mean precipitation. Manuel liked the suggestion to provide a guide on best practice and the hourly data as well as a daily and monthly mean created by us following the guide of best practice. In addition, we have iterated the issue with re-gridding the data. ECMWF provides tools for users to re-grid their data to their needs, however, Phil Jones will prepare also a guide on best practice on re-gridding the users should follow. WP2 and WP3 will procedures will follow that guide.

Data can be archived with hardly any restriction on size. However, ECMWF would like to archive only data products which users are likely to use. Preliminary estimates on storage size from the producers:

Eric Bazile	Peter Jermey	Jan Keller	Per Undén
50 years (MESCAN and SURFEX) and 5 years MESCAN-Ensemble 1) For internal expert users 40 TB 2) For external users (hourly output with analysis field and SURFEX output - ~16 variables) 35TB Grand Total: max 100TB	Regional Deterministic - analysis/forecasts hourly and three-hourly forecasts to T+30 twice daily --------------------------------------- Surface fields 12.6TB Pressure level fields 11.0TB Model level fields 7.85TB Total 31.5TB Regional Ensemble - analysis/forecasts hourly; 20 members --------------------------------------- Surface fields 36.2TB Pressure level fields 31.7TB Model level fields 22.6TB Total 90.5TB Grand Total 122TB	10 members COSMO for 5 years: 30 TB	50 years @ 11km - HARMONIE archiving requirement: 600 TB A bit depending on how much of the fields should be archived of course, but it should be quite inclusive.

14. Reference list

Bowler, N. E., 2006: Explicitly accounting for observation error in categorial verification. doi: 10.1175/MWR3138.1

E.J.M. van den Besselar, A.M.G. Klein Tank, G. van der Schrier, and P.D. Jones, 2011: Synoptic messages to extent climate data records. doi: 10.1029/2011JD016687

Candille, G., and O. Talagrand, 2008: Impact of observational error on the validation of ensemble prediction systems. doi: 10.1002/qj.221

Ferro and Stephenson, 2011: Extremal Dependence Indices: Improved Verification Measures for Deterministic Forecasts of Rare Binary Events. doi: 10.1175/WAF-D-10-05030.1

J. Qiu, W.T. Crow, G.S. Nearing, X. Mo, and S. Liu, 2014: The impact of vertical measurement depth on the information content of soil moisture time series data. doi: 10.1002/2014GL060017