Seventh Framework Programme

Theme 6 [SPACE]

Project: 607193 UERRA

Full project title:

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

Deliverable D3.2

Preliminary table summarizing common evaluation procedures shared among WP3 partners

WP no:	3
WP leader:	DWD
Lead beneficiary for deliverable :	DWD
Name of author/contributors:	Michael Borsche, Andrea Kaiser-Weiss with UERRA partners
Nature:	Report
Dissemination level:	PU
Deliverable month:	6
Submission date: October 2, 2014	D3.2-v20140930

UERRA D3.2: Preliminary table summarizing common evaluation procedures shared among WP3 partners

1. Preliminary table summarizing common evaluation procedures

Method	Data source	Parameter	Details	Scientific questions	User questions
A: feedback statistics	Radiosonde soundings	Temperature wind speed relative humidity	Focus on lower troposphere; bias and RMSE of time series; store in ODB format;	How stable are the regional reanalyses (RRAs) with respect to multi-annual trends on a spatial scale of roughly 100 km?	How well represented are trends and climatologies of wind speed at heights which are relevant for wind energy?
B: point measurements	B1: (independent) mast station data; B2: (dependent, i.e., assimilated) station data	B1: wind speed B2: Tmin, Tmax, and number of days of threshold exceedance of temperature and precipitation	There are many more suitable observations available for B2 than for B1.	At which time scales can we find correlations between reanalysis fields and station observations?	On which time scales of variability can we use the RRAs (for which parameters) similar to the use of a station measurement?
C: gridded measurements	Gridded data products for the Nordic region and the UK; EOBS, APGD	Precipitation; Tmin and Tmax	To consider whether a part of underlying station observations was assimilated into the reanalysis.	What differences do we get with different products when determining the useful spatial and temporal scales of the RRAs?	Which scales of the RRAs (temporal, spatial) can be interpreted?

D: satellite data products	Satellite data products of CM-SAF and CCI	Global radiation; total cloud cover; snow water equivalent		How well do the RRAs fit to the satellite observations - or exceed their quality?	Depending on the parameter, is the RRA or the satellite the better data product for the user to use?
E: Ensemble based comparison	WP1 created ensemble of gridded data with derived uncertainty estimates; products as in methods A through D	precipitation , Tmin, Tmax, Tmean; parameters as in A through D	Ensemble based uncertainty estimates will be performed on (1) the newly (WP1) created data products. (2) the basis of methods A through D when available.	Does the ensemble provide a more detailed and spatially and temporal more resolved estimate of uncertainty compared to a deterministic reanalysis?	Which uncertainty characteristics can be interpreted from the reanalyses ensembles for user relevant parameters?
F: User related models		Tmean; Tmax and Tmin pseudo analysis; wind speed; precipitation;	SURFEX by Météo France uses the reanalyses as input		Is the result of a user model forced by RRAs significantly better than with the original forcing?

2. Discussion on evaluation methods

The following discussion is an excerpt of the protocol of the WP3 workshop (D3.1) held at Deutscher Wetterdienst, Offenbach, Germany on 26/27 June 2014. This discussion was the basis of the results presented in the above table.

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. 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, on observatory including a large mast in The Netherlands. There should be other such stations which measure at tall masts in Europe. Andrea Kaiser-Weiss, Michael Borsche, and Frank Kaspar have checked and confirmed that independent mast measurements are available for Germany from Deutscher Wetterdienst.

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 base 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 Jones and Christoph Frei have written a guide on best practices on aggregation and regridding, which can be read in Appendix A1. Further clarification from Eric Bazile and Cornel Soci concluded that in the Météo France reanalysis product Tmin and Tmax will be calculated as a pseudo analysis based on the 2m temperature 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 evaluate 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 Met Office 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 Met Office. 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. Comparing the forecast to the analysis of each system and calculating system specific diagnostics from this might be misleading when comparing the different systems. The difference between forecast and analysis is not an objective measure of the quality of the analysis.

On the other hand, we agreed on a detailed plan on which re-forecasts we would like to store for subsequent use in WP3 for precipitation uncertainty evaluation as discussed in methods B and C, as well as for the general users.

3. 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 1. 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].

Appendix

A1. Guide on Aggregation of data and Regridding

This guide was written by Phil Jones affiliated at Climate Research Unit, University of East Anglia, Norwich, UK and Christoph Frei affiliated at MeteoSwiss, Zurich, Switzerland in August 2014.

Background

All National Met Services (NMSs) across Europe use not only different ways of archiving the basic station meteorological data, but also employ different approaches to combining the measurements of the key variables into daily and monthly averages. The issues that this causes are highlighted by van den Besselaar et al. (2012) when attempting to use the SYNOP network to update daily NMS temperature and precipitation series. The 24-hour period to which the 'day' refers to differs between NMSs as does the time the measurements were made. Also some record the measurements against the time the instrument was read, while some assign the readings (such as maximum temperature or the rainfall accumulation) to the period when it occurred (i.e. the day before).

To improve real-time updating capabilities, GCOS has asked for NMSs to produce a daily CLIMAT message and to issue this at the end of the month (all days in the month together) when the monthly CLIMAT message is sent. It will take some time for this request to be acted upon, if it is.

Implications

The impacts of these issues are different for different variables. The simplest variable is precipitation. If precipitation amounts are accessed/used from SYNOP data, the totals will be for the two 12-hour periods from 06-18 and 18-06. This total will be similar (but not that same) for NMSs that use a day definition such as 06-06 or 07-07, but will differ more for an NMS that has traditionally used the period 09-09 for the day (such as the UK).

Monthly precipitation totals produced in CLIMAT reports are rounded to the nearest whole mm. Similarly daily precipitation totals in SYNOP reports are also rounded to whole mm, except for values below 0.9mm where they are in units of 0.1mm. Validated data (see van den Besselaar et al., 2012) received directly from NMSs will be in mm and tenths.

Validated maximum and minimum temperatures supplied by NMSs will refer to the definition of the day used by that NMS.

Aggregation of Data

Reanalysis output provides daily maximum and minimum temperatures and precipitation totals. For comparisons with observed data (either validated from NMSs or from the SYNOP network) values for the 24-hour period 06-06 should be used, as these will best approximate what will have been observed. It is important for Reanalysis daily temperatures that the two extreme temperatures also cover the full 24-hour period from 06-06. If the Regional Reanalysis (RRA) is running at additional times to every 6 hours, then it would be possible to accumulate rainfall totals over say 09-09 for the UK.

Aggregation of RRA output to the daily time step needs to be mindful of the definition of the day used by the NMSs for the study area. It is likely that a compromise will have to be sought except for specific national studies.

Regridding

It is very likely that the resolution of RRA products within UERRA will differ from the various gridded daily datasets that are available for validation assessments. Possible products include European-wide datasets such as E-OBS, and more highly resolved regional products such as the Alpine dataset (updated recently by Isotta et al., 2014) or national datasets (such as that for the UK at 5 by 5 km, Perry and Hollis, 2005 and Perry et al., 2008).

Most of these potential validation products are constructed with the intent that a grid-point value represents area mean conditions across the grid box. In many cases, the representativity is however coarser, i.e. grid-point values represent mean conditions over many grid boxes. Reasons for this are the coarse density of stations and the principle of "best estimates" employed in gridding. As a result, grid values in a dataset of gridded station data should be interpreted as averages over one to many grid boxes. The area of averaging is mostly not known and it may vary spatially as the station density varies. The scale discrepancy between areal and point values is significant. For precipitation it means that grid-box values will tend to have more frequent raindays and that high quantiles are smaller than in individual station series (e.g. Gervais et al. 2014). The fact that the value for each grid box is an average across an area is more important for precipitation series than for temperature, but the same principles apply to temperature so in regions of strong relief the average value will more likely relate to the average elevation of the grid box.

Discrepancies in areal representativity are also significant between grid datasets and RRA, because the averaging area may be quite different between them. Disregarding these discrepancies during evaluation may unduly degrade the apparent skill of a RRA.

In comparing RRA with grid datasets it is essential to ensure compatibility in averaging scales. This is generally best achieved by regridding the finer to the coarser grid. A common approach is the aggregation (spatial averaging) of all grid points of the fine grid within the boxes of the coarser grid. However, this may not eliminate scale discrepancies sufficiently when the coarser grid dataset has an areal representativity coarser than its grid resolution (with RRA this is to be expected). It may then be appropriate to upscale to even coarser scales. For parameters with a strong elevation dependence (notably temperature), the regridding (upscaling) needs to make sure that test and reference datasets have comparable elevation. For parameters with strong spatial variance in the climatological mean distribution (e.g. temperature with its elevation dependence), regridding is best undertaken in anomaly space with the mean field (the climatology at a monthly scale) and deviations regridded separately and then added back.

References

Gervais, M., L. B. Tremblay, J. R. Gyakum, and E. Atallah, 2014: Representing Extremes in a Daily Gridded Precipitation Analysis over the United States: Impacts of Station Density, Resolution, and Gridding Methods. *J. Clim.*, **27**, 5201–5218, doi:10.1175/JCLI-D-13-00319.1.

Isotta, F.A. et al., 2014: 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**, 1657-1675.

Perry, M. and Hollis, D., 2005: The generation of monthly gridded datasets for a range of climate variables across the UK. *International Journal of Climatology* **25**: 1041-1054.

Perry, M., Hollis, D. and Elms, M., 2008: The generation of daily gridded datasets for temperature and precipitation, covering the UK for the period 1960 to 2006. Internal report, Met Office, Exeter, UK.

van den Besselaar, E.J.M. Klein Tank, A.M.G, van der Schrier, G. and Jones, P.D., 2012: Synoptic messages to extend climate data records. *Journal of Geophysical Research*, **117**, D07101, doi:10.1029/2011JD1688.

A2. Parameter Table

This preliminary parameter table has been set up to provide an overview on the output specifics of each regional reanalysis from HErZ, Météo France, Met Office, and SMHI. This table serves as a means to agree on common levels and parameters to store. Levels are divided into model, pressure, and surface levels. There is also the idea to provide a certain set of parameters on interpolated height levels. So far, the parameters to be output by each reanalysis producer have been collected and summarized. WP3 has indicated by blue highlighting which parameters need to be output in order to perform the verification task as outlined in the workpackage.

Model Levels

Parameter	Met Office		HErZ		SMHI	
	Analysis	Forecast	Analysis	Forecast	Analysis	Forecast
Cloud cover	X	X	X	X	X	X
Cloud ice	X	X	X	X	X	X
Cloud liquid water content	X	X	X	X	X	X
Potential temperature	X	X				
Pressure	X	X	X	X		
Specific humidity	X	X	X	X	X	X
Specific rain water content			X	X	X	X
Specific snow water content			X	X	X	X
Temperature			X	X	X	X
U component of wind	X	X	X	X	X	X
V component of wind	X	X	X	X	X	X
Vertical velocity			X	X		

Model levels to output (approximate height and pressure values)

	Met Office			HErZ			SMHI	
Level	Height [m]	Pressure [hPa]	Level	Height [m]	Pressure [hPa]	Level	Height [m]	Pressure [hPa]
1	10	1000	35	258.5	969.73	56	313.2	979.28
2	37	997	36	178.5	979.02	57	272.4	983.89
3	77	992	37	116.0	986.32	58	200.4	988.15
4	130	986	38	69.0	991.84	59	168.4	992.11
5	197	978	39	34.5	995.91	60	138.7	995.81
			40	10.0	998.81	61	111.0	999.27
						62	84.6	1002.56
						63	59.8	1005.70
						64	35.6	1008.74
						65	11.8	1011.75

Met Office:

- Model levels follow terrain and flatten towards the top of the model following Charney-Phillips staggering
- Model level values are interpolated in log(pressure) to provide fields on requested pressure surfaces

HErZ:

- Approximate heights and pressure are only valid for grid points above sea
- with orography these values can (dramatically) change due to the orography following coordinate

SMHI:

- relative to the model grid elevations / terrain following
- pressure relative to 1013.25 hPa surface pressure
- heights are approximate using a surface temperature of 273K

Pressure levels

Parameter	Met Office		HErZ		SMHI	
	Analysis	Forecast	Analysis	Forecast	Analysis	Forecast
Cloud cover			X	X	X	X
Geopotential height	X	X	X	X	X	X
Potential vorticity						X
Relative humidity	X	X	X	X	X	X
Specific cloud ice water content			X	X		X
Specific cloud liquid water content			X	X		X
Specific humidity			X	X	X	X
Temperature	X	X	X	X	X	X
U component of wind	X	X	X	X	X	X
V component of wind	X	X	X	X	X	X
Vertical velocity			X	X		X

Pressure levels to output

Met Office [hPa]	HErZ [hPa]	SMHI [hPa]	Standard pressure levels [hPa]
1000	1000	1000	1000
850	975	950	925
700	950	925	850
500	925	900	700
300	900	850	500
200	875	800	400
100	850	700	300
	825	600	250
	800	500	200

	750	400	150
	700	300	100
	600	250	
	500	200	
	400	150	
	300	100	
	250		
	200		
	150		
	100		

Standard pressure levels go up to 100 hPa. The pressure levels above (which corresponds to the climatological tropical tropopause), are not of primary interest in connection with the regional reanalyses. The standard levels have been taken from the following WMO definition:

WMO, 1996: Guide to Meteorological Instruments and Methods of Observation. 6th ed. WMO Rep. 8, World Meteorological Organization, Geneva, Switzerland

Surface/1D Parameters

Parameter	Météo France		Met Office		HErZ		SMHI	
	Analysis	Forecast	Analysis	Forecast	Analysis	Forecast	Analysis	Forecast
1.5 m specific humidity	X	X	X	X	X	X	X	X
10 m U wind [over land | component]	X	X	X	X	X	X	X	X
10 m V wind [over land | component]	X	X	X	X	X	X	X	X
10 m wind gust [in the last 24 hrs | since previous post-processing]		X	X	X				
1.5/2 m temperature	X	X	X	X	X	X	X	X
2 m dewpoint temperature			X	X	X	X	X	X
Accumulated Total Precipitation	X	X	X	X	X	X		

Parameter	Météo France Analysis	Météo France Forecast	Met Office Analysis	Met Office Forecast	HErZ Analysis	HErZ Forecast	SMHI Analysis	SMHI Forecast
between 6TU and 6[TU] day+1								
Albedo	X	X			X	X	X	X
Clear-sky (II) down surface sw flux			X	X				X
Clear-sky (II) up surface sw flux			X	X				X
Convective available potential energy					X	X	X	X
Convective precipitation		X	X	X	X	X		X
Convective rain rate					X	X		X
Convective snowfall		X	X	X	X	X		X
Convective snowfall rate					X	X		X
Evaporation	X	X			X	X		X
Geopotential height	X	X	X	X	X	X	X	X
High cloud cover		X			X	X		X
Instantaneous surface sensible heat flux								X
Large-scale rainfall rate					X	X		X
Large-scale snowfall rate					X	X		X
Large-scale precipitation			X	X	X	X		X
Long-wave radiation flux					X	X		X
Low cloud cover		X			X	X	X	X
Maximum [1.5 m | 2 m] temperature since previous post-processing			X	X	X	X	X	X
Mean sea level pressure	X	X	X	X	X	X	X	X
Medium cloud cover		X			X	X	X	X
Minimum [1.5 M | 2 m] temperature since previous post-processing			X	X	X	X	X	X
Net long-wave radiation flux (surface)	X	X			X	X		X
Net short-wave radiation flux (surface)	X	X	X	X	X	X		X
Sea surface temperature					X	X	X	X
Fraction of sea-ice in sea			X	X				
Skin temperature	X	X	X	X	X	X		

Parameter	Météo France Analysis	Météo France Forecast	Met Office Analysis	Met Office Forecast	HErZ Analysis	HErZ Forecast	SMHI Analysis	SMHI Forecast
Snow albedo	X	X					X	X
Snow density	X	X			X	X	X	X
Snow depth	X	X			X	X	X	X
Soil temperature level 1	X not the same level	X not the same level	X	X	X	X	X	X
Soil temperature level 2	X not the same level	X not the same level	X	X	X	X	X	X
Soil temperature level 3	X not the same level	X not the same level	X	X	X	X	X	X
Soil temperature level 4			X	X	X	X		
Surface latent heat flux	X mean flux?	X mean flux?	X	X	X	X		X
Surface net solar radiation	X	X			X	X		X
Surface net thermal radiation	X	X			X	X		X
Surface pressure	X	X	X	X	X	X	X	X
Surface roughness	X	X	X	X				
Surface sensible heat flux	X mean flux?	X mean flux?	X	X	X	X		X
Surface solar radiation dwn (flux)		X	(X)	(X)				
Surface thermal radiation downwards		X						
TOA incident solar radiation								X
Temperature of snow layer		X 3 layers			X	X		
Top net solar radiation (flux)			(X)	(X)	X	X		
Total cloud amount – random overlap			X	X				
Total cloud amount in lw radiation			X	X				
Total cloud cover		X	X	X	X	X	X	X
Total column water vapour			X	X	X	X	X	X
Total precipitation	X	X	X	X	X	X		X
Visibility at 1.5 m			X	X				
Volumetric soil water layer 1	X not the	X not the	X	X	X	X	X	X

Parameter	Météo France Analysis	Météo France Forecast	Met Office Analysis	Met Office Forecast	HErZ Analysis	HErZ Forecast	SMHI Analysis	SMHI Forecast
	same level	same level						
Volumetric soil water layer 2	X not the same level	X not the same level	X	X	X	X	X	X
Volumetric soil water layer 3	X not the same level	X not the same level	X	X	X	X	X	X
Volumetric soil water layer 4	X not the same level	X not the same level	X	X	X	X		
Water equivalent of accumulated snow depth	X	X	X	X	X	X		

Static fields

For fields that don't change (much or at all), e.g., land mask, orography, etc., we should store them separately, to save space.

Parameter	ERUO4M	ERA-Interim	Météo France		Met Office		HErZ		SMHI	
			Analysis	Forecast	Analysis	Forecast	Analysis	Forecast	Analysis	Forecast
Land cover (1=land, 0=sea)	X		X	X	X	X			X	X
Orography			X	X	X	X	X	X	X	X
Forecast surface roughness		X			[X]	[X]				X

A3. Comparison Table

This comparison table was initiated at the WP3 workshop D3.1. It summarizes essential characteristics of each regional reanalysis participating in the UERRA project. It is a living document and therefore subject to change.

Feature	HErZ	HErZ	Met Office	SMHI	Météo France
Boundary conditions (forcings)	3 hourly ERA-Interim analyses and forecasts	3 hourly ERA-Interim and/or ERA-20C fields	6 hourly ERA-Interim fields	6 hourly ERA-Interim fields	HARMONIE
Model	COSMO	COSMO	Unified Model	HARMONIE	MESCAN forced by a downscaled (11km to >5.5km) HARMONIE (SMHI) or by ALADIN at 5.5km
Grid (projection) and Domain	CORDEX EU-11	CORDEX EU-11	CORDEX EU-11	Lambert, Europe-Atlantic 11km, as CORDEX-EU	Lambert
Ensemble members	1	10-20	20	1 (2 for a period)	1
DA method	Continuous nudging; Kalman filter analyses for soil moisture; Interpolation methods for SST, sea ice, and snow cover	Hybrid Ensemble Nudging / Ensemble Kalman Filter	4D Ensemble Variational	3D Variational upper-air / OI surface analysis	OI surface re-analysis after a static or dynamical downscaling
Short-term forecast	forecast basetime and steps	forecast basetime and steps	forecast basetime and steps	forecast basetime and steps	forecast basetime and steps
Time range	1978 to present	5 years	1978 to present	1961 to 2011	1961 to 2011

Observation input	SYNOP (pressure), SHIP, BUOY, DRIBU, AIREP, AMDAR, TEMP, PILOT, T2M for SM	SYNOP (pressure), SHIP, BUOY, DRIBU, AIREP, AMDAR, TEMP, PILOT	Surface (land SYNOP, METAR, SHIP, BUOY, E-Obs gridded precipitation, ground GPS); upper air (TEMP, drop sondes, pilot balloons, wind profilers); aircraft (AIREPS, AMDAR); satellite (SSMI/S, AMV, AIRS, (A)TOVS, IASI,GPSRO,scat winds, clear sky radiances)	SYNOP (pressure), SHIP, BUOY, DRIBU, AIREP, AMDAR, TEMP, PILOT	SYNOP (pressure), SHIP, BUOY, 24h precipitation from rain gauge and Tmin/Tmax after pre-processing
Temporal resolution	1 hour 3D, 15 minutes 2D	1 hour	6 hours (analysis), hourly (forecast)	6 hours	6 hours (or 3 hours) except for precipitation 24h
Horizontal resolution	6km	12 km	12km grid; analysis increments on 24km; 24km ensemble (expected)	11 km	5.5km
Vertical resolution	40 levels (20m to 22km)	40 levels (20m to 22km)	70 levels from near surface to 80km	65 levels	only surface
Main reference	Bollmeyer et al., 2014, in revision		Rawlins et al., 2007, The Met Office global four-dimensional variational data assimilation scheme. Q.J.R. Meteorol. Soc., 133: 347\u2013362. doi: 10.1002/qj.32		http://www.euro4m.eu/downloads/D2.6_Report_describing_the_new_system_in_D2.5.pdf