ECMWF COPERNICUS NOTE

Copernicus Climate Change Service

UERRA data user guide

Issued by: SMHI / S. Schimanke

Date: 17/01/2020

Ref: C3S_322_Lot1.4.1.2_UERRA_data_user_guide – version 3.3

Official reference number service contract: 2017/C3S_322_Lot1_SMHI/SC1

IMPLEMENTED BY ECMWF

Contributors

SMHI
S. Schimanke
L. Isaksson
L. Edvinsson
P. Undén
M. Ridal

MÉTEO-FRANCE
P. Le Moigne
E. Bazile
A. Verrelle
M. Glinton

1. Introduction

1.1 The service

The C3S_322_Lot1 part of the Copernicus Climate Change Service produces and delivers a regional reanalysis (RRA) for Europe including long-term datasets of Essential Climate Variables (ECVs). These datasets can be used in support of adaptation action and policy development as well as contribute to climate services, climate monitoring and research.

The service is implemented in several steps. First, a system developed in the FP7 pre-operational project UERRA (UERRA: Uncertainties in Ensembles of Regional Reanalyses; www.uerra.eu) was used to update the existing RRA in near real time. In combination with the RRA produced already in the pre-operational project, the service offers a consistent RRA from January 1961 – July 2019. The production of the UERRA RRA ceased with the end of availability of ERA-interim.

Moreover, an improved model version was developed within the service. The model will be used to create a pan-European reanalysis with very high resolution (5.5 km) forced by the global ERA5 reanalysis (RA). The improved system started the production at the beginning of 2020.

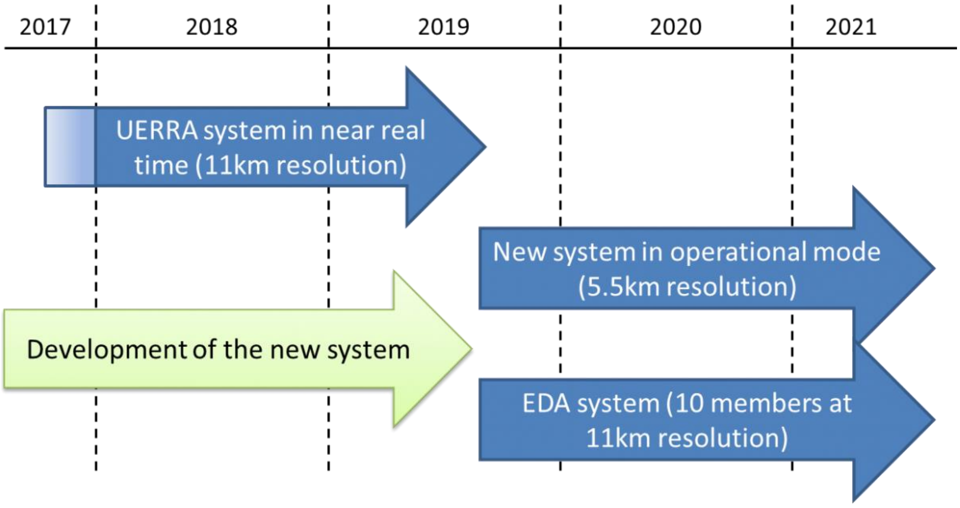


Figure 1: The scheduled time line of production with the different systems.

This document aims to describe the data produced by the first model system (UERRA system), which was developed in the pre-operational UERRA project. The system consists of two parts: the 3-dimensional model version called UERRA-HARMONIE at 11 km resolution and the 2-dimensional MESCAN-SURFEX surface analysis at 5.5 km resolution.

The UERRA-HARMONIE Data Assimilation system is implemented and optimized for the European area with surrounding sea areas (see Figure 2) with a resolution of 11 km and 65 height levels. It was run for the period January 1961 – July 2019. The production ceased with the end of availability of ERA-interim data.

The MESCAN-SURFEX 2D-analysis system is used to generate a surface analysis. The system combines downscaled UERRA-HARMONIE reanalysis fields and additional surface observations to make a 2-dimensional analysis with 5.5km resolution over Europe for the period January 1961 – July 2019.

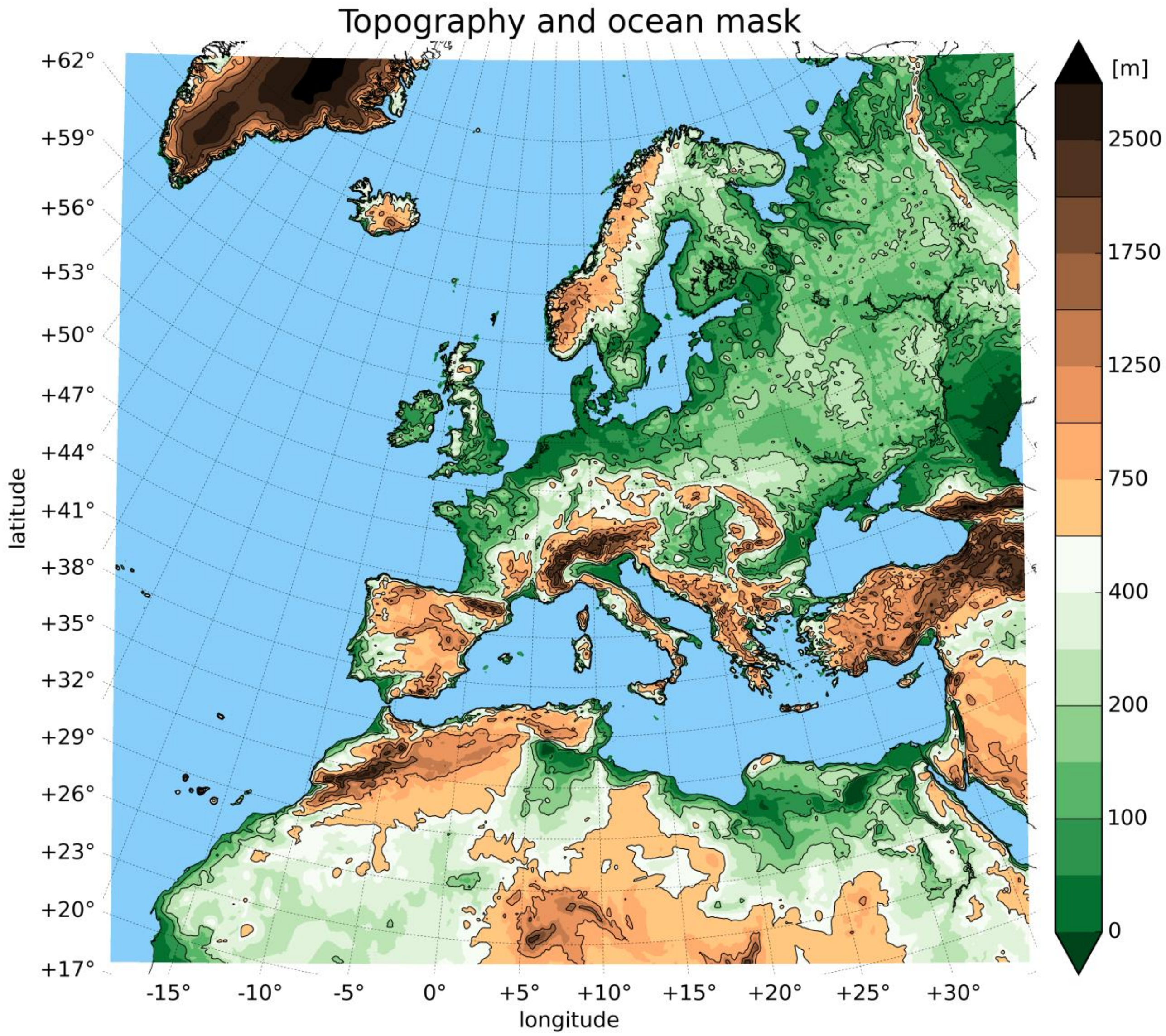


Figure 2: Model domain of the UERRA systems highlighted with the topography of the UERRA-HARMONIE system where ocean cells are masked in blue.

1.2 Principles of reanalysis systems

Atmospheric reanalysis is a method to reconstruct the past weather by combining historical observations (in situ, surface and satellite remote sensing), with a dynamical model. It provides a physically and dynamically coherent description of the state of the atmosphere. The synthesis is accomplished by assimilating the observational data into a meteorological model and thereby forcing the model to reproduce the observations as closely as possible. The advantage of reanalysis is that they provide a multivariate, spatially complete, and coherent record of the atmospheric state – far more complete than any observational dataset is able to achieve.

The main advantages of reanalyses are (see Verver 2017):

- They provide regularly gridded data, even in places where there are no or few observations;
- They provide a coherent, complete set of variables describing the atmospheric state;
- They provide a reconstruction of the record of past weather since it is constrained by observations.

Weather forecasting is based on an analysis of the current state of the atmosphere and the surface of land and sea. The forecasts are made with mathematical and physical computer models starting from the analysis. The temperatures, winds, pressure, moisture, cloud contents and other variables are mapped at regular points in space and time (Fig. 3).

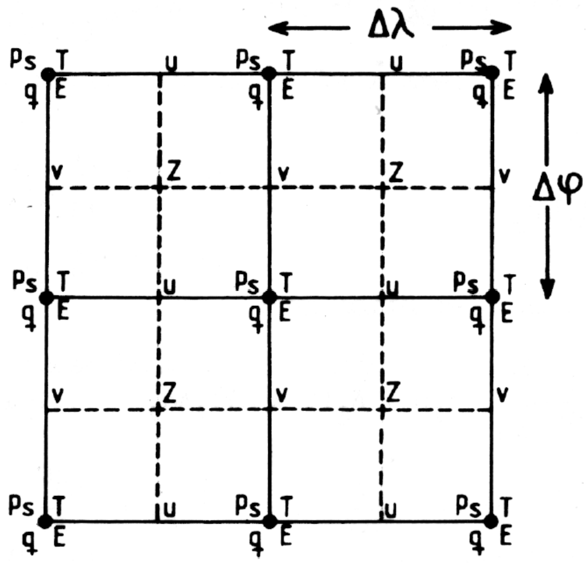


Figure 3: Schematic representation of a grid for model variables (surface pressure, temperature, u and v wind components and geopotential (Z), Energy and specific moisture content (q)).

Reanalysis uses a weather forecasting model to create a 'first guess' of the atmospheric state at a certain time. The first guess is then corrected on the basis of observations. This corrective step, referred to as 'data assimilation' (see figure 4), requires statistical knowledge of the forecast error and the observation error. The procedure also uses physical and statistical relationships of the atmosphere when interpreting the observational data. The result of the data assimilation is called the analysis. By repeating this process for a number of time steps the analyses will contain a complete set of values describing the evolution of the atmosphere and the surface over time, also for locations where there are no observations.
This complete estimate of the atmospheric state over time can be of great value to users, for example in assessing impacts of past weather and climate related events, for statistics of the climate in a location or an area or for running other fine scale models or validating climate models.

An important difference between reanalyses and archived weather analyses from operational forecasting systems is that a reanalysis is produced with a single version of a data assimilation system – including the forecast model used – and is therefore not affected by changes in method.

Reanalysis systems differ in the set of observations that are assimilated, the model that is used, and the way the error statistics are estimated and corrections are applied. A variety of reanalysis methods exist as for instance 4D variational analysis (4D-VAR), 3D-VAR (schematically shown in Fig. 4), nudging, and optimal interpolation (OI).

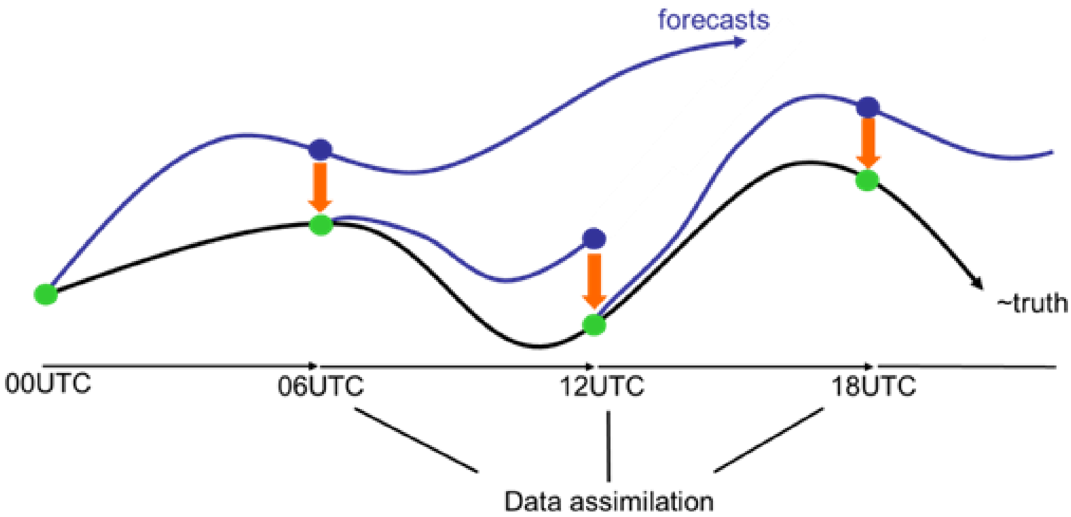


Figure 4: Schematic showing the simulation of the atmospheric state (black line) in the reanalysis, which starts from the analysis (green dots) and resulting in the background (blue dots). Note that the background usually does not coincide with the true observed state of the atmosphere. The source of the figure is unknown.

1.3 The UERRA system

As stated above, the model system that is used during the first phase of the service was designed in the preoperational project UERRA. That is why it is called UERRA system even in the framework of this service. Here, we will briefly mention the main features of the system whereas a detailed technical description can be found in Ridal et al. (2018).
In general, the UERRA system consists of two components. The three dimensional reanalysis based on the UERRA-HARMONIE system as well as the surface reanalysis based on MESCAN-SURFEX.

1.3.1 The UERRA-HARMONIE system

The 3D-UERRA system is based on the HARMONIE Data Assimilation system, which is developed and used within the HIRLAM and ALADIN consortia. The UERRA system is implemented and optimized for the entire European area with surrounding sea areas (see Fig. 2) with a resolution of 11 km and 65 height levels. For the period 1961-2001 ERA40 observations with addition of Swedish and French observations are used. After 2001 conventional data (SYNOP, Ship, Buoys, Radiosondes, Pilot and Aircraft) are used that are operationally available.
The system uses global reanalysis data as lateral boundaries: ERA40 for the period 1961-1978, after that ERA-interim (Fig. 5). Also, the large scales in the regional system are constrained by data from the global reanalysis.

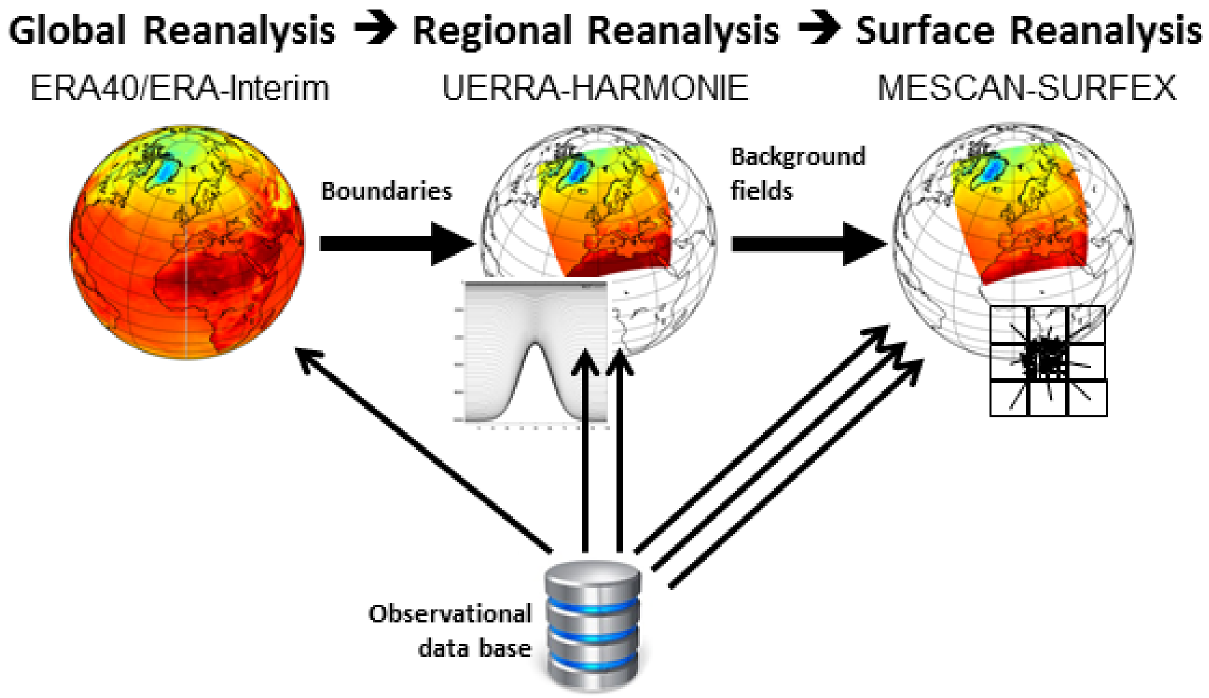


Figure 5: Three different stages of RA: the global reanalyses are used as lateral boundary for the 3D-UERRA regional reanalysis. In turn, these are used as background fields for the 2D surface reanalysis. The amount of observational data used for the RA per area unit increases from the global to the surface reanalysis as indicated by the arrows.

The UERRA system applies the so called 3D variational analysis (3D-VAR). The 3D-VAR method is depicted schematically in Figure 4. At fixed points in time the model state is adjusted based on the observed state, taking into account the statistics of model and observation errors. The UERRA system is run with four cycles per day performing analyses at 00UTC, 06UTC, 12UTC and 18UTC. The forecasts lengths vary between 6 and 30 hours.

1.3.2 MESCAN-SURFEX

The MESCAN-SURFEX system analysis uses the 2D-analysis system MESCAN and the land surface platform SURFEX to generate a coherent surface and soil analysis. The system combines downscaled UERRA-HARMONIE reanalysis fields and additional surface observation (especially for precipitation), to make a high resolution (5.5 km) 2-dimensional analysis over Europe.
MESCAN is a surface analysis system using an optimal interpolation algorithm for the 2m-temperature and relative humidity and for the 24h-accumulated rainfall (Soci et al., 2016).
SURFEX is a land surface platform, which is driven by temperature, humidity, precipitation, wind and radiative fluxes.

More details regarding the MESCAN-SURFEX system can be found in Bazile et al. (2017).

2. General guidelines for the usage of UERRA data

This section aims to summarize important features of the models and results that the user needs to be aware of when using the UERRA datasets. Although UERRA provides consistent and coherent datasets there are weaknesses and limitations. Some of these are common for reanalyses in general, other are model/version dependent. The user has to decide whether the data is fit for their specific purpose.

Vast amount of data are available from the UERRA system. For the UERRA-HARMONIE system, a complete set of parameters including all available time steps take up ca. 8 TB per model year and hence almost 500 TB for the entire time period. The MESCAN-SURFEX surface analysis is not included in this estimation. However, it needs less storage place than the 3D-reanalysis – in total 47 TB.

For the UERRA-HARMONIE data, the user might choose among more than 50 parameters for different heights and time steps. In addition, the MESCAN-SURFEX analysis offers roughly 30 surface parameters. Detailed descriptions of all parameters are given in section 3.

2.1 Resolution in time and space

2.1.1 Horizontal resolution

As depicted in Fig. 3, all parameters are computed for grid boxes. This means the parameters values reflect a mean over the grid box area. Having a horizontal resolution of 11km for the UERRA-

HARMONIE system implies that each value reflects the mean over an area of 121km^2 (11km*11km). This needs to be considered when for instance UERRA data is compared with observation.
The resolution of the MESCAN-SURFEX surface analysis is 5.5km*5.5km. Hence, a grid box has an area of roughly 30km^2.

2.1.2 Vertical resolution

As stated above, the UERRA-HARMONIE system has 65 vertical levels. However, on these so called model levels only a very restricted amount of parameters is stored. The main reason for that is the amount of needed storage space, when all parameters would be stored for all levels. Moreover, the vertical model grid is on hybrid-sigma coordinates, which makes it more complex to use the output.

Therefore, the major part of the data is stored on selected pressure levels. Pressure levels are available at levels between 1000-10hPa with a higher resolution at lower altitudes.
In addition, some parameters are stored on height levels. 11 height levels are available which are between 15-500m. One reason to store data additionally on height levels are applications in the wind energy sector.
The exact levels both for pressure and height levels are given in section 3 in the corresponding tables.

The UERRA-HARMONIE soil model has 3 vertical levels. The three levels represent approximately the surface, the soil at root depth and the deep soil. Due to the used force-restore scheme in the soil model it is not possible to relate the levels with a certain depth in meter.
The MESCAN-SURFEX soil model has 14 vertical levels, which range from the surface to a depth of 12m. The edges between different levels are at 0.01m, 0.04m, 0.1m, 0.2m, 0.4m, 0.6m, 0.8m, 1.0m, 1.5m, 2m, 3m, 5m, 8m, and 12m. Values for a certain level reflect the mean value over the level thickness.

2.1.3 Time resolution

In general, data is stored with hourly resolution for the UERRA-HARMONIE system. However, for many time steps the users have different options to select from and this is no easy choice. The preferred selection might vary for different parameters and the application of the user, respectively. Also, some of the time steps are affected by spin-up issues as explained in the next section.
Figure 6 gives an overview on available time steps. First, there are the four analyses at 00UTC, 06UTC, 12UTC, and 18UTC highlighted in red. These time steps should be of highest quality since the observations are assimilated directly. However, they are available only every sixth hour and not all parameters are available for the analyses (Check tables in section 3 for the availability of the parameters at different time steps).
The forecasts are then started from the analyses and the output is saved hourly for the first six hours as indicated by the dark blue in Fig. 6. Whereas the forecasts initiated at 06UTC and 18UTC stop after six hours the forecasts initiated at 00UTC and 12UTC continue until forecast hour 30. However, the output frequency is reduced to three hourly until forecast hour 24 and the last output is then saved six hours later (see blue boxes in Fig. 6).

Due to the forecast lengths, the forecasts are overlapping and for many hours of the day data might be chosen from different forecasts and the analysis, respectively. At 00UTC and 12 UTC, the users can choose between the analysis and three different forecasts. As shown in Fig. 6, at 12 UTC

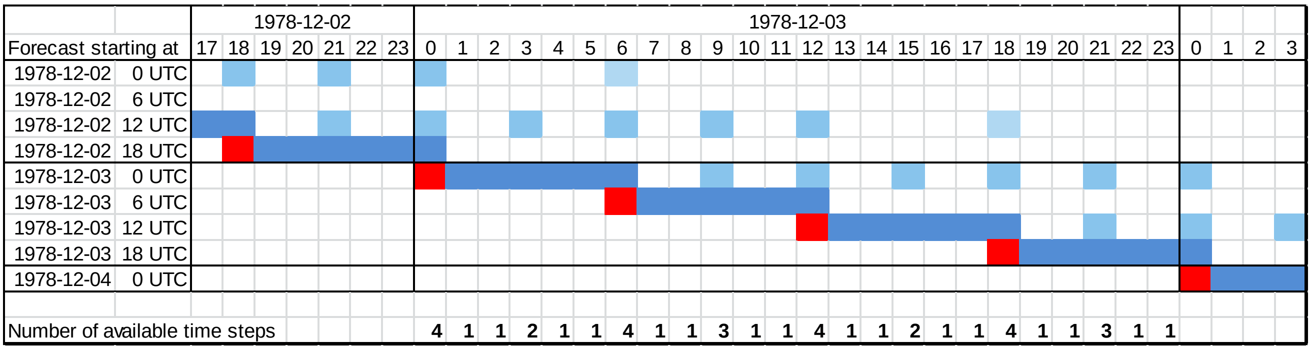


Figure 6: Availability of time steps in the UERRA-HARMONIE system illustrated for the example date 1978/12/03. The color coding reflects analysis (red) and forecasts (blue). Moreover, different shades of blue correspond to frequency of the saved forecasts – hourly in dark blue and 3/6hourly forecasts in blue.

results are available from the forecasts initiated at 06UTC (six hour forecast), 00UTC (twelve hour forecast), and 12 UTC of the previous day (24 hour forecast). However, for other hours of the day (01UTC, 02 UTC, 04 UTC, …) there is no choice and results are available only from one forecast. Different forecasts lengths have different strengths and weaknesses. Whereas the short-term forecasts are affected by spin-up issues after the initialization of the model the long-term forecasts might veer away from the real weather due to shortcomings in the model. In general, it is not possible to give a general recommendation for which time steps should be used and the users have to check on their own, which selection gives the best result in their application.

The MESCAN-SURFEX output is essentially hourly except for the driving variables used as input by SURFEX. The analyzed variables are 2m temperature and relative humidity, radiative fluxes and wind with a frequency of 6h and 24h precipitation only available at 6h UTC.

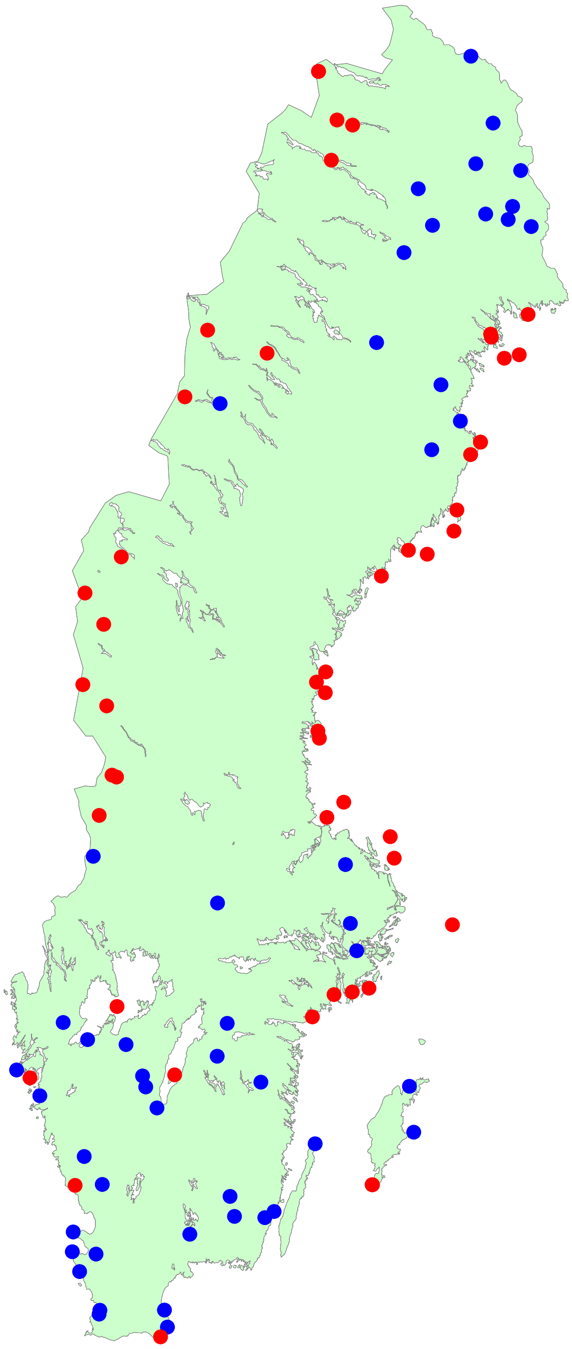
2.2 General limitations of reanalyses

Generally it is challenging for a reanalysis system to correctly reconstruct variables that is very variable in space and time, such as precipitation. For some applications, e.g. in hydrology, it is therefore quite common to correct the precipitation data for a bias. Other variables, like surface temperature, are generally less variable in space and time and easier to reconstruct by the reanalysis system.

Similar to above, results in complex terrain, such as mountainous regions or coastal areas, are generally less reliable than results over a more homogeneous terrain. The models cannot represent the strong gradients that sometimes are caused by the variable terrain.
Figure 7 illustrates this behavior. Here, we show spots in Sweden having the best (blue) and worst (red) correlations between the UERRA-HARMONIE 2m-temperature and observational sites. A total

of 853 measurement sites have been investigated and each the 50 with highest and lowest correlation are shown. Clearly, correlations are lowest in the Swedish mountains and along the (east) coast.

Users need to keep in mind, that the reanalysis provides gridded data where each grid box describes the mean over the grid box area. That's in contrast to observations, which are usually point measurements. In case users need information with a higher horizontal resolution than provided by the UERRA-systems, further downscaling (statistically or dynamically) needs to be considered. For instance, the correlations indicated in Fig. 7 increased when a linear interpolation to the observational site was applied than purely taken the values from the closest grid point.



Partially due to this it is more difficult for a reanalysis system to correctly capture absolute values of extremes than values closer to the mean. This is especially the case for precipitation extremes, where the reanalysis data are highly resolution dependent. This means for example that the number of days with precipitation over a certain absolute threshold value is likely to be less accurate than using a relative threshold such as a 95-percentile value. Also extremes on larger scales, like droughts and heatwaves, are better represented than extremes on smaller scales.

As mentioned earlier, the reanalysis is produced with a single version of data assimilation system/forecast model and is therefore not affected by changes in method. But it is worth noting that some other components of the reanalysis are not consistent over time. For example, the number of available observations is varying over time. Also, the shift of global reanalysis boundary data in 1980 and observation data sources in 2001 will affect the consistency of the time series (see section 1.3.1).

Figure 7: A validation of the UERRA-HARMONIE 2m-temperature data with Swedish observations. 50 places each with highest (blue) and lowest (red) correlation are shown out of 853 measurement sites included in the investigation.

2.3 Known problems

In the UERRA reanalyses some problems have been noticed that users need to be aware of. As mentioned above, the amount of data is very large and only part of the data was verified. Hence, often problems are only discovered by the end users. To increase the performance of future reanalysis, we would appreciate if users report known problems. This section lists the issues that are known at the time of writing. However, this is a living document, which is updated if new issues are discovered. In case of severe issues are discovered users who downloaded data via CDS might be informed via email.

2.3.1 UERRA-HARMONIE

2.3.1.1 Spin-up issues

Spin-up issues are a general problem for data assimilation and Numerical Weather Prediction systems. However, unfortunately this particular issue is in the UERRA-HARMONIE case somewhat extreme. But, the issue causing the problem is located and the next RRA should not be affected as much by spin-up problems as the current version.

The largest spin-up issues occurs during the first 2 hours of the forecast. It is therefore recommended to primarily use the 6 hourly analysis fields whenever this is possible. Also, it seems that three hourly values can be used without larger restriction. The spin-up problems are most pronounced during turbulent conditions, for example connected to deep low pressure systems, and should therefore be limited in space and time.

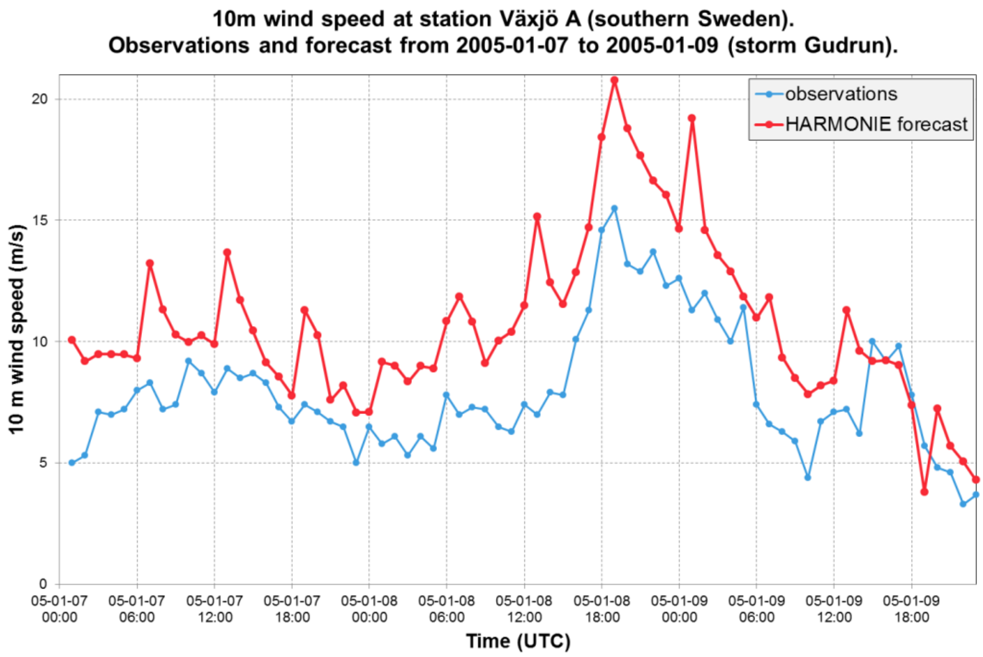


Figure 8: 10m wind speed during the storm Gudrun in southern Sweden January 2005. The blue line and dots are hourly observations from station Växjo A. Red line and dots are the hourly UERRA-HARMONIE forecast for the closest grid point. Note the reoccurring leaps in wind speed during forecast hours 1-2.

These spin-up problems are in some situations causing too high wind speeds for the 1-2 hour forecasts. This is most pronounced in the wind gust but the problem also occurs for the mean wind at 10 meter altitude, see figure 8.

Similar spin-up problems as for the winds can be seen for maximum/minimum temperature. This can in turn cause errors in the cold/warm extremes which affect climate indices such as frost days, tropical nights, ice days and summer days. For instance, Niermann et al. (2017) report that biases for frost days and summer days can be as large as up to 40 days/year when compared with E-OBS data.

Then in general there is a spin-up problem for precipitation. Here, it is not recommended to use the first 6 hours of the forecasts. The spin-up of the precipitation is illustrated in Fig. 9, which shows the precipitation for different forecast lengths aggregated for the entire model domain. It is quite clear that the precipitation stabilized after forecast hour 6 – beside the daily cycle during summer. Therefore, if a user wants to use the 24 hour accumulated precipitation it is recommended to use the 30-6 hour forecasts.

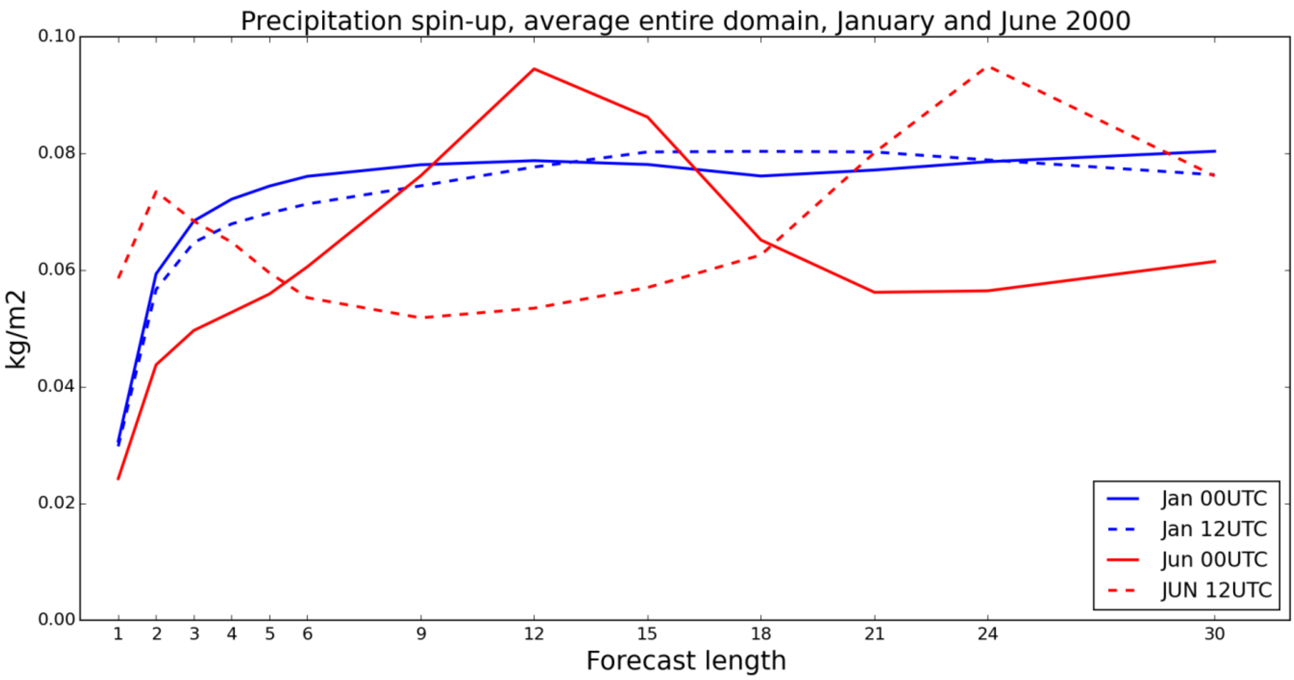


Figure 9: Average precipitation in relation to the forecasts length for January and June 2000.

2.3.1.2 Archiving issues – relative humidity, cloud cover and albedo

Relative humidity: Due to a bug in one of the archiving scripts all relative humidity equal or larger than 100% was set to zero. That issue can be handled by users by setting all zero values to 100%. At least close to the surface there is no risk of setting real zeros to 100 since the humidity is never reaching 0% - not in the model and not in the real world.

For users familiar with CDOs (climate data operators, https://code.mpimet.mpg.de/projects/cdo) this can be done with the following command line:

cdo setmisstoc,100 –setvrange,0.0001,200 INFILE OUTFILE

Cloud cover: The same problem exists for cloud cover so there are large areas archived with 0% that are erroneous. The problem exists for total cloud cover as well as for low, medium and high clouds. Unfortunately, this problem cannot be solved as easily as for humidity since the cloud cover can be zero, which is not the case for relative humidity (in the lower atmosphere). There might be ways to fix this parameter as well. A combination of radiation parameters might be checked to verify whether the cloud cover is really zero or should be set to 100% instead. However, such a reprocessing has not been tested yet.

Albedo: The same issue can be probably found in albedo fields as well but its extent there is not known.

The archiving bug was fixed within the Copernicus service and values saved from RRA year 2016 are not affected by this problem.

2.3.1.3 Surface fluxes

For the accumulated surface fluxes of sensible and latent heat values are missing over most of the land area. Whereas the model handles the values correctly, they were ruined in the post-processing. The problem is understood and will be fixed for future production.

And, the user is encouraged to use the surface fluxes from the MESCAN-SURFEX system.

2.3.1.4 Radiation/Clouds

For global radiation, there is a general overestimation over Europe. The bias is largest over the ocean parts of the North Atlantic and decreases towards the south. Details regarding global radiation can be found in the UERRA report by Niermann et al. (2017) in section 5.2.

It is suspected that the overestimation of global radiation can be a consequence of underestimated cloud cover and investigations are ongoing.

2.3.1.5 Relative humidity/Clouds

Ridal et al. (2018) states that UERRA-HARMONIE produces too much moisture and that this is a known problem for the ALADIN model. The excess of moisture does affect the cloud cover. For instance, the UERRA-HARMONIE re-analysis performs worse than ERA-Interim in situations with low amounts of clouds.

2.3.1.6 Geopotential height

For geopotential height, it seems that some time steps are erroneous. The values seem wrong by a factor of 10.000. The reason for this sporadic error is unknown. Users reported errors for the following days: 1st and 2nd of January 1981, 1st – 4th of August 1981, and 11th of February 1982.

2.3.2 MESCAN-SURFEX

2.3.2.1 2m temperature

Due to biased observational data and a too tolerant quality control, analyzed temperatures can differ severely from other most likely more realistic analysis, e.g. UERRA-HARMONIE or E-OBS. Differences can be more than 3K for the annual mean as noticed for Poland in 1965 and much more for instantaneous values. **Users are requested to perform additional sanity checks when using this parameter or to switch to the T2m of the UERRA-HARMONIE system.**

More details on the problem:
Issues with the quality of 2-meter air temperature (T2m) data were discovered, namely "potential" cold biases particularly evident in 1965 in Poland and the UK along the coast, and also a warm bias located in 2010 in the Baltic countries. However, other regions and periods might be affected as well. The issues were initially found from a comparison between the MESCAN analyses, UERRA-HARMONIE analyses and the gridded E-OBS dataset constructed from an interpolation of observations. And, the issue becomes also clear when comparing with the first guess as done in Fig. 10 for the year 1965.

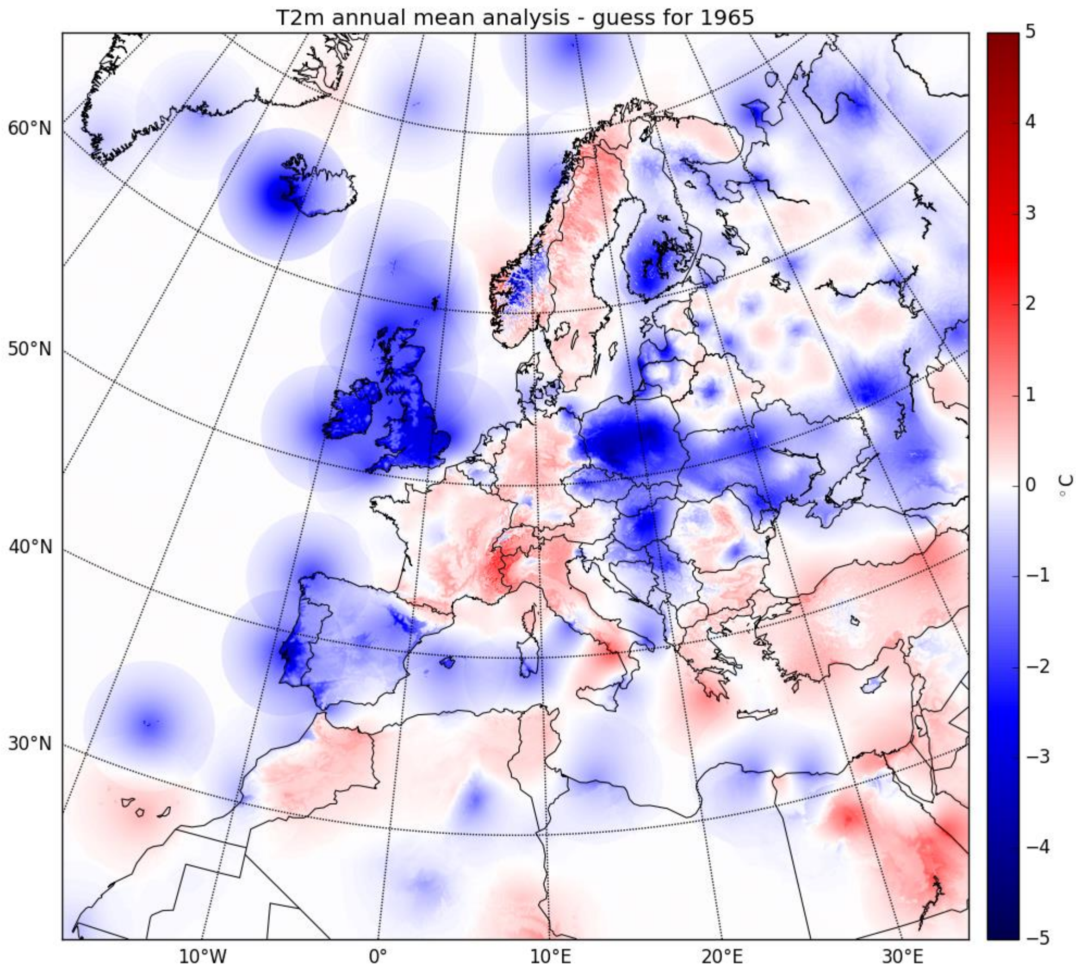


Figure 10: First guess minus analysis averaged for the entire year 1965.

A more detailed inspection of the cold bias in Poland in July 1965 showed that some stations, only 3, with systematic cold bias of about 5-10 degrees Celsius, were identified as originating from the

MARS input data stream. For example, in July 1965 only those 3 stations are available in Poland instead of 20-40 and these 3 stations have the same bias.
Another aspect is the automatic quality control based on the background (6-hour model forecast) whose parameters used to determine the threshold value for rejection have been adjusted in order to have an analysis closer to the observations. This tuning used for reanalysis is very beneficial in case of an observation input database being well validated. However, it now seems that the automatic quality control of the T2m observation must be stricter to reject observations when they deviate too much from the model (as it is in numerical weather prediction in real time).

2.3.2.2 Precipitation
In the MESCAN-SURFEX surface analysis errors have been noticed in the precipitation fields for some areas and time periods, see Figure 11. These errors are due to erroneous observations, which did enter the assimilation procedure. It remains a challenge for the next system to solve this type of weaknesses – especially for “old” and validated data where precipitation is equal to zero for long period.
Since the errors are restricted in time and space**, the user has to evaluate if the studied time period or region is affected**.

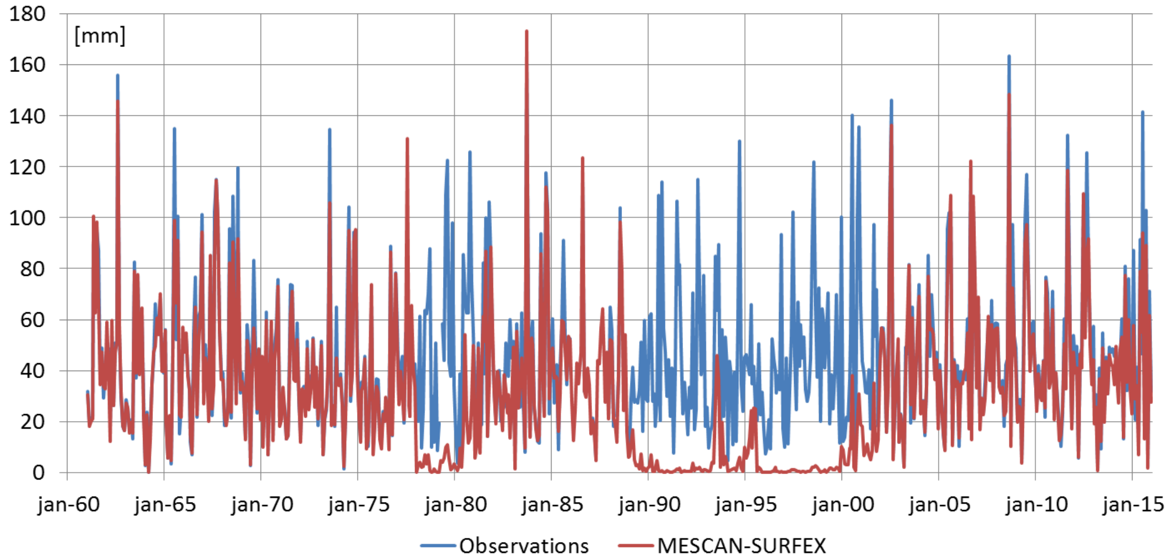


Figure 11: Monthly precipitation at station Enköping/Sweden (blue line) and from the closest grid point of the MESCAN-SURFEX surface analysis (red line). Note the two periods where precipitation is underestimated in MESCAN-SURFEX.

2.3.2.3 Time-integrated surface solar radiation downwards
The time-integrate surface solar radiation downwards which comprises of both the direct and diffuse short wave radiation reaching the Earth's surface, ought to be positive by definition. Whilst archiving, it was found that it also contains negative values that were traced back to being an artefact from the downscaling from 11 km to 5.5 km. As a work around, we would suggest setting any negative values to zero.

2.3.2.4 Land-sea mask

The SURFEX-MESCAN system uses two different land-sea masks: the ALADIN5.5km mask when dynamically downscaling from 11km to 5.5km and the SURFEX mask when making the analysis. The SURFEX mask is available as a downloadable parameter, but the ALADIN5.5km mask is not included in the archived output. Since this land-sea mask might be of interest for users it can be found here: https://confluence.ecmwf.int/display/UER/Issues+with+data#

The SURFEX land-sea mask has been noticed to only contains values for land (=1) and water (=0), whereas the system handles gridboxes internally with a mixture of different surfaces, e.g. partly land and partly water. This means that the SURFEX land-sea mask can be 1 even if the land fraction is below 10%.

Because of its higher resolution the ALADIN5.5km land-sea mask introduces new lakes that were not resolved in the 11km UERRA-HARMONIE system. This difference causes problems for the model which makes it fall back to using climatological values for these new lake areas.

These issues regarding the MESCAN-SURFEX land-sea masks can cause unrealistic results for coastal areas and lake affected grid points. It is therefore not recommended to use data over lakes.

2.3.2.5 Soil level parameters are inverted with respect to their Y-axis

Both soil level fields, Volumetric transpiration stress-onset (soil moisture) and Volumetric wilting point, are inverted with respect to their Y-axis.

2.4 FAQ

- What do you mean with near real-time?
  The production of the UERRA system is delayed by 3-4 month. For instance, in May 2018 we released the RRA for Europe for January 2018.
  The delay is directly coupled to the production and validation of the global RA ERA-interim, which is used as lateral boundary. Only if this data is released are we able to run the UERRA-HARMONIE system and then MESCAN-SURFEX in a second step.

- Can we use reanalysis data for local applications?
  Reanalysis data are gridded products. The values represent a certain spatial scale which may be hard to compare to a point value that may be obtained from a station. Note that the spatial scale of the data provided is not necessarily the grid spacing of the dataset.

- What's recommended for the computation of daily/monthly means? Which time steps should be considered?
  Some tests were performed at SMHI regarding 2m-temperature. In one case, the monthly mean was computed from hourly values – the analysis and the first 5 forecasts time steps. In the second case, only the analyses were considered. Consequently each day was represented by 4 values only whereas in case one 24 hourly values were used. However, the mean bias of the monthly means compared to observations was generally somewhat lower when only analyses were used.
  Users should keep in mind that this test was done for 2m-temperature only and that other parameters might behave differently.

- Precipitation and snow is represented in the model output, but not rain?
  The rain can be calculated as the amount of precipitation that is not snow, total precipitation minus the precipitation as snow.

- Is the data free and how can it be accessed?
  The UERRA reanalysis datasets (in total more than 700 Tb) are freely available at the Copernicus Climate Data Store (CDS) at https://cds.climate.copernicus.eu/ . The full set of variables that are available can be found in https://software.ecmwf.int/wiki/display/UER/Parameters. Details on how to retrieve data from this archive can be found in section 4 of this document.

- What is the projection used? Is there an EPSG code?
  The used grids are in the Lambert Conformal Conic projection with parameters according to section 3.1.6 for UERRA-HARMONIE and section 3.2.3 for MESCAN-SURFEX. There are no defined EPSG codes for these specific projections, but there is a general definition of Lambert Conformal Conic with 2 standard parallels in EPSG:9802.

  Unfortunately, the coordinate systems are often problematic when handling model output. But a great improvement is that nowadays many available programs are able to handle grib files and especially the projection information that is embedded in the UERRA grib files. For example later versions of many GIS tools (QGIS, ArcGIS) can open and georeference the grib files out of the box. Just drag and drop a grib file into the GIS tool and it should open up, see Fig. 12. This can be a good starting point for getting to know the grid and handle the projection.

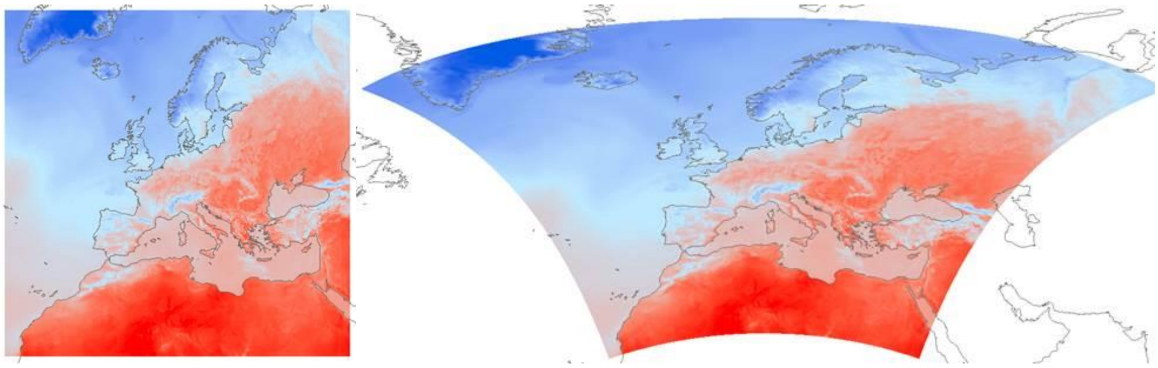


Figure 12: A MESCAN-SURFEX grib file opened in QGIS. To the left in the models Lambert Conformal Conic projection, to the right transformed to WGS 84.

QGIS also generates proj strings which for example can be used to transform coordinates with the widely used proj library (https://proj.org/). Below are the QGIS (v. 2.8.6) generated proj strings:

UERRA-HARMONIE proj string:
+proj=lcc +lat_1=48 +lat_2=48 +lat_0=48 +lon_0=8 +x_0=0 +y_0=0 +a=6371229 +b=6371229 +units=m +no_defs

MESCAN-SURFEX proj string:
+proj=lcc +lat_1=50 +lat_2=50 +lat_0=50 +lon_0=8 +x_0=0 +y_0=0 +a=6371229 +b=6371229 +units=m +no_defs

Worth noting here is that these generated proj strings set the false easting and false northing parameters (x_0 and y_0) to zero. These can be used to assure that the coordinates are positive in the Lambert Conformal Conic system. These should not influence the georeferencing of the data, but they can be useful in other applications. One way of calculating these parameters is by using the cdo program. Below are the false easting and false northing values calculated by cdo (v. 1.9.7.1):

UERRA-HARMONIE:
false easting = 3101971.52403412
false northing = 2994156.36404286

MESCAN-SURFEX:
false easting = 2937018.5829291
false northing = 2937031.41074803

The mentioned cdo program is excellent for many grib handling tasks. But please note that before version 1.9.7.1 it did not include support for the earth radius used in the UERRA grib files. Previous versions of cdo report an earth radius of 6367470 m instead of the correct 6371229 m.

One possible use of the false easting/northing could be transforming between grid cell index and WGS 84 coordinates. This can for example be performed with the proj command cs2cs. Below are examples of such transformation commands:

UERRA-HARMONIE, grid cell index (zero based) to WGS 84 coordinates:
cs2cs -f %.5f +proj=lcc +lat_1=48 +lat_2=48 +lat_0=48 +lon_0=8 +x_0=3101971.52403412 +y_0=2994156.36404286 +a=6371229 +b=6371229 +to_meter=11000 +to +proj=lonlat +ellps=WGS84 +datum=WGS84

UERRA-HARMONIE, WGS 84 coordinates to grid cell index (zero based):
cs2cs -f %.0f +proj=lonlat +ellps=WGS84 +datum=WGS84 +to +proj=lcc +lat_1=48 +lat_2=48 +lat_0=48 +lon_0=8 +x_0=3101971.52403412 +y_0=2994156.36404286 +a=6371229 +b=6371229 +to_meter=11000

MESCAN-SURFEX, grid cell index (zero based) to WGS 84 coordinates:
cs2cs -f %.5f +proj=lcc +lat_1=50 +lat_2=50 +lat_0=50 +lon_0=8 +x_0=2937018.5829291 +y_0=2937031.41074803 +a=6371229 +b=6371229 +to_meter=5500 +to +proj=lonlat +ellps=WGS84 +datum=WGS84

MESCAN-SURFEX, WGS 84 coordinates to grid cell index (zero based):
cs2cs -f %.0f +proj=lonlat +ellps=WGS84 +datum=WGS84 +to +proj=lcc +lat_1=50 +lat_2=50 +lat_0=50 +lon_0=8 +x_0=2937018.5829291 +y_0=2937031.41074803 +a=6371229 +b=6371229 +to_meter=5500

3. Detailed data description and availability

3.1 UERRA-HARMONIE system

3.1.1 Surface parameters

Metadata for UERRA-HARMONIE surface parameters	
Horizontal coverage	The model domain spans from northern Africa beyond the northern tip of Scandinavia. In the west it ranges far into the Atlantic ocean and in the east it reaches to the Ural. Herewith, it covers entire Europe. See Figure for an overview of the model domain.
Horizontal resolution	11x11 km^2.
Vertical coverage	Each surface parameter is valid for one vertical level and three different (near) surface levels exists: - surface: atmospheric boundary with the ground or water surface - 2m: 2m above the surface - 10m: 10m above the surface
Vertical resolution	One level only
Temporal coverage	1961-01-01 00:00 – close to real time (monthly updates but with a delay to real time of about four months)
Temporal resolution	Analysis are available at 00, 06, 12, 18 UTC. The forecast length is depending on the cycle. Cycles initialized at 00 and 12 UTC have forecasts saved at 1, 2, 3, 4, 5, 6, 9, 12, 15, 18, 21, 24, 30. Cycles initialized at 06 and 18 UTC have forecasts saved at 1, 2, 3, 4, 5, 6. Some parameters are saved only until forecast hour six. See section 2.1.3 (Fig. 6) for more details.
Data type and format	Gridded data in GRIB2
Grid	Lambert conformal conic grid with 565x565 grid points

Table 1: Overview Surface parameters

	Name	Unit	GRIB code	Analysis 0, 6, 12, 18	Forecast 1,2,3,…	Height
1.	2m relative humidity	%	260242	yes	yes	2m
2.	Total column integrated water vapour	kg/m^2	260057	yes	yes	vertically integrated above the surface
3.	Total precipitation	kg/m^2	228228	-	yes	surface
4.	10m wind speed	m/s	207	yes	yes	10m

5.	10m wind direction	degrees	260260	yes	yes	10m
6.	10m wind gust speed	m/s	49	-	yes	10m
7.	Surface air maximum temperature	K	201	-	yes	2m
8.	Surface air minimum temperature	K	202	-	yes	2m
9.	2m temperature	K	167	yes	yes	2m
10.	Skin temperature	K	235	yes	yes	surface
11.	Albedo	%	260509	yes	only six hours	surface
12.	Evaporation	kg/m^2	260259	-	yes	surface
13.	Time-integrated surface latent heat flux	J/m^2	147	-	only six hours	surface
14.	Time-integrated surface sensible heat flux	J/m^2	146	-	only six hours	surface
15.	Time-integrated surface direct solar radiation	J/m^2	260264	-	yes	surface
16.	Time-integrated surface net solar radiation	J/m^2	176	-	yes	surface
17.	Time-integrated surface solar radiation downwards	J/m^2	169	-	yes	surface
18.	Time-integrated surface net thermal radiation	J/m^2	177	-	yes	surface
19.	Time-integrated surface thermal radiation downwards	J/m^2	175	-	yes	surface
20.	Mean sea level pressure	Pa	151	yes	yes	surface
21.	Surface pressure	Pa	134	yes	yes	surface
22.	High cloud cover	%	3075	yes	yes	above 5000m
23.	Low cloud cover	%	3073	yes	yes	surface-2500m
24.	Medium cloud cover	%	3074	yes	yes	2500m-5000m
25.	Total cloud cover	%	228164	yes	yes	above ground
26.	Snow density	kg/m^3	33	yes	only six hours	surface
27.	Snow depth water equivalent	kg/m^2	228141	yes	only six hours	surface
28.	Snow fall water equivalent	kg/m^2	228144	-	yes	surface
29.	Land-sea mask	dimension-less	172	yes	-	surface
30.	Orography	gpm	228002	yes	-	surface
31.	Surface roughness	m	173	yes	only six hours	surface

3.1.1.1 2m relative humidity

The 2m relative humidity (parameter name surface air relative humidity) is the modelled humidity valid for a grid point (approximately 11km*11km=121km^2) determined for a height of 2m above the surface. The parameter is the relation between actual humidity and saturation humidity given in %. Values are in the interval [0,100]. 0% means that the air is totally dry whereas 100% indicates that the air is saturated with water vapour. The saturation is defined with respect to saturation of the mixed phase, i.e. with respect to saturation over ice below -23°C and with respect to saturation over water above 0°C. In the regime in between a quadratic interpolation is applied.
Surface air relative humidity is available for the analysis and the forecast time steps. For the forecast, the value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.
Please check section 2.3.1.2 Archiving issues when using this parameter.

3.1.1.2 Total column integrated water vapour

The total column integrated water vapour is the vertically integrated water vapour valid for a grid point (approximately 11km*11km=121km^2). It is vertically integrated from the surface to the top of the atmosphere. The parameter is given in kg/m^2.
Total column water vapour is available for the analysis and the forecast time steps. For the forecast, the value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.1.3 Total precipitation

Total precipitation is the amount of precipitation falling onto the ground/water surface. It includes all kind of precipitation forms as convective precipitation, large scale precipitation, liquid and solid precipitation. The amount is valid for a grid box and has the unit kg/m^2.
The total precipitation is available only for the forecast time steps. It is an accumulated parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 24h-forecast has the accumulated precipitation over 24 hours.
The total amount of rain might be computed by subtracting the snowfall water equivalent.
Please be aware that precipitation is affected from the spinning up of the model. Check section 2.3.1.1 for more details.

3.1.1.4 10m wind speed

The 10m wind speed is the wind speed valid for a grid point (approximately 11km*11km=121km^2) determined for a height of 10m above the surface. The parameter is given in m/s. It is computed from both the zonal (u) and the meridional (v) wind components by

$$wind\ speed = \sqrt{u^2 + v^2}$$

The 10 metre wind speed is available for the analysis and the forecast time steps. For the forecast, the value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.1.5 10m wind direction

The 10m wind direction is the wind direction valid for a grid point (approximately 11km*11km=121km^2) determined for a height of 10m above the surface. The parameter is given in degrees ranging from 0-360. Here, 0° means a northerly wind and 90° indicates an easterly wind.
The 10 metre wind direction is available for the analysis and the forecast time steps. For the forecast, the value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.1.6 10m wind gust speed

The 10 metre wind gust speed is the maximum wind speed since the last post-processing at a grid point (approximately 11km*11km=121km^2). It is determined for a height of 10m above the surface. The parameter is given in m/s.

The 10 metre wind speed is only available for the forecast time steps. The value is the maximum since the previous post-processing. For instance, for the first saved time step at forecast 1h it is the maximum wind speed, which occurred between within the first hour of the forecast. For the second saved time step at forecast 2h, it is the maximum wind speed which happened in the second forecast hour, hence between fc1 and fc2. For longer forecasts, the output frequency is reduced. Hence, the maximum over a longer time period is saved. For instance, for the 15h forecast the maximum wind speed is identified within the period 12h – 15h since the last post-processing happened at 12h (12 hours after the onset of the forecast).

Unfortunately, there is a quality problem for this parameter. Please check section 2.3 Model specific issues when using this parameter.

3.1.1.7 Surface air maximum temperature

The surface air maximum temperature is the maximum temperature since the last post-processing at a grid point (approximately 11km*11km=121km^2). It is determined for a height of 2m above the surface. The parameter is given in Kelvin [K].

Surface air maximum temperature is only available for the forecast time steps. The value is the maximum since the previous post-processing. For instance, for the first saved time step at forecast 1h it is the maximum surface air temperature, which occurred between within the first hour of the forecast. For the second saved time step at forecast 2h, it is the maximum surface air temperature which happened in the second forecast hour, hence between fc1 and fc2. For longer forecasts, the output frequency is reduced. Hence, the maximum over a longer time period is saved. For instance, for the 15h forecast the maximum surface air temperature is identified within the period 12h – 15h since the last post-processing happened at 12h (12 hours after the onset of the forecast).

Please check section 2.3 Model specific issues when using this parameter.

3.1.1.8 Surface air minimum temperature

The surface air minimum temperature is the minimum temperature since the last post-processing at a grid point (approximately 11km*11km=121km^2). It is determined for a height of 2m above the surface. The parameter is given in Kelvin [K].

Surface air minimum temperature is only available for the forecast time steps. The value is the minimum since the previous post-processing. For instance, for the first saved time step at forecast 1h it is the minimum surface air temperature, which occurred between within the first hour of the forecast. For the second saved time step at forecast 2h, it is the minimum surface air temperature which happened in the second forecast hour, hence between fc1 and fc2. For longer forecasts, the output frequency is reduced. Hence, the minimum over a longer time period is saved. For instance, for the 15h forecast the minimum surface air temperature is identified within the period 12h – 15h since the last post-processing happened at 12h (12 hours after the onset of the forecast).

Please check section 2.3 Model specific issues when using this parameter.

3.1.1.9 2m temperature

The 2m temperature with the parameter name surface air temperature is the model temperature valid for a grid point (approximately 11km*11km=121km^2) determined for a height of 2m above the surface. The parameter is given in Kelvin [K].

Surface air temperature is available for the analysis and the forecast time steps. For the forecast, the value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.1.10 Skin temperature

The skin temperature is the model temperature valid for a grid point (approximately 11km*11km=121km^2) determined for the boundary surface to the atmosphere, both ground and water surfaces. The parameter is given in Kelvin [K].

Skin temperature is available for the analysis and the forecast time steps. For the forecast, the value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.1.11 Albedo

The albedo is the amount of radiation which is reflected for the given grid point (approximately 11km*11km=121km^2). It is determined for the surface to the atmosphere, both for ground and water surfaces. The parameter is given in %. Small values mean that large amounts of the radiation are absorbed whereas large values mean that more radiation is reflected.

Albedo is available for the analysis and the forecast time steps up to forecast hour six. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

Please check section 2.3 Model specific issues when using this parameter.

3.1.1.12 Evaporation

Evaporation is the amount of moisture flux from the surface (ground and water) into the atmosphere. It is given as a mean for a grid point (approximately 11km*11km=121km^2). The parameter is given in kg/m^2. By model convention downward fluxes are positive. Hence, evaporation is represented by negative values and positive values represent condensation.

Evaporation is only available for forecast time steps. It is an accumulated parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 24h-forecast has the accumulated evaporation over 24 hours.

3.1.1.13 Time-integrated surface latent heat flux

The surface latent heat flux is the exchange of latent heat (due to phase transitions: evaporation, condensation) with the surface (ground and water) through turbulent diffusion. It is given as a mean for a grid point (approximately 11km*11km=121km^2). The parameter is given in J/m^2. By model convention downward fluxes are positive.

Surface latent heat flux is only available for forecast time steps up to forecast hour six. It is an accumulated (time-integrated) parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 24h-forecast has the accumulated heat flux over 24 hours.

Please check section 2.3 Model specific issues when using this parameter.

3.1.1.14 Time-integrated surface sensible heat flux

The surface sensible heat flux is the exchange of heat (no phase transition) with the surface (ground and water) through turbulent diffusion. It is given as a mean for a grid point (approximately 11km*11km=121km^2). The parameter is given in J/m^2. By model convention downward fluxes are positive.

Surface sensible heat flux is only available for forecast time steps up to forecast hour six. It is an accumulated (time-integrated) parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 24h-forecast has the accumulated heat flux over 24 hours.

Please check section 2.3 Model specific issues when using this parameter.

3.1.1.15 Time-integrated surface direct solar radiation
The surface direct solar radiation is the amount of direct solar (short-wave) radiation reaching the surface (ground and water). It is given as a mean for a grid point (approximately 11km*11km=121km^2). The parameter is given in J/m^2. By model convention downward fluxes are positive.
Surface direct solar radiation is only available for forecast time steps. It is an accumulated (time-integrated) parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 24h-forecast has the accumulated radiation over 24 hours.
3.1.1.16 Time-integrated surface net solar radiation
The surface net solar radiation is the amount of solar (short-wave) radiation that is absorbed at the surface (ground and water). It is computed

*Surface net solar radiation = surface solar radiation downwards * (1 – albedo)*

It is given as a mean for a grid point (approximately 11km*11km=121km^2). The parameter is given in J/m^2. By model convention downward fluxes are positive.
Surface net solar radiation is only available for forecast time steps. It is an accumulated (time-integrated) parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 24h-forecast has the accumulated radiation over 24 hours.
3.1.1.17 Time-integrated surface solar radiation downwards
The surface solar radiation downward is the amount of solar (short-wave) radiation reaching the surface (ground and water). It is given as a mean for a grid point (approximately 11km*11km=121km^2). The parameter is given in J/m^2. By model convention downward fluxes are positive.
Surface solar radiation downwards is only available for forecast time steps. It is an accumulated (time-integrated) parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 24h-forecast has the accumulated radiation over 24 hours.
3.1.1.18 Time-integrated surface net thermal radiation
The surface net thermal radiation is the amount of thermal (long-wave) radiation that is absorbed at the surface (ground and water). It is computed

*Surface net thermal radiation = surface thermal radiation downwards * (1 – albedo)*

It is given as a mean for a grid point (approximately 11km*11km=121km^2). The parameter is given in J/m^2. By model convention downward fluxes are positive.
Surface net thermal radiation is only available for forecast time steps. It is an accumulated (time-integrated) parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 24h-forecast has the accumulated radiation over 24 hours.
3.1.1.19 Time-integrated surface thermal radiation downwards
The surface thermal radiation downward is the amount of thermal (long-wave) radiation reaching the surface (ground and water). It is given as a mean for a grid point (approximately 11km*11km=121km^2). The parameter is given in J/m^2. By model convention downward fluxes are positive.
Surface thermal radiation downwards is only available for forecast time steps. It is an accumulated (time-integrated) parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 24h-forecast has the accumulated radiation over 24 hours.
3.1.1.20 Mean sea level pressure
The mean sea level pressure is the air pressure reduced to mean sea level valid for a grid point (approximately 11km*11km=121km^2). The parameter is given in Pascal [Pa].

Mean sea level pressure is available for the analysis and the forecast time steps. For the forecast, the value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.1.21 Surface pressure

The surface pressure is the air pressure at the surface (ground and water) valid for a grid point (approximately 11km*11km=121km^2). The parameter is given in Pascal [Pa].

Mean sea level pressure is available for the analysis and the forecast time steps. For the forecast, the value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.1.22 High cloud cover

The high cloud cover is the percentage of sky covert with clouds in high altitude. It is valid for a grid point (approximately 11km*11km=121km^2) and high refers to height above 5000m. The parameter is given in %.

High cloud cover is available for the analysis and the forecast time steps. For the forecast, the value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

Please note that there is a problem with this parameter. Check section 2.3.1.2 Archiving issues for more details when using this parameter.

3.1.1.23 Low cloud cover

The low cloud cover is the percentage of sky covert with clouds in low altitude. It is valid for a grid point (approximately 11km*11km=121km^2) and low altitude refers to heights below 2500m. The parameter is given in %.

Low cloud cover is available for the analysis and the forecast time steps. For the forecast, the value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

Please note that there is a problem with this parameter. Check section 2.3.1.2 Archiving issues for more details when using this parameter.

3.1.1.24 Medium cloud cover

The medium cloud cover is the percentage of sky covert with clouds in medium altitude. It is valid for a grid point (approximately 11km*11km=121km^2) and medium altitude refers to heights between 2500m through 5000m. The parameter is given in %.

Medium cloud cover is available for the analysis and the forecast time steps. For the forecast, the value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

Please note that there is a problem with this parameter. Check section 2.3.1.2 Archiving issues for more details when using this parameter.

3.1.1.25 Total cloud cover

Total cloud cover is the percentage of sky covert with clouds. It is valid for a grid point (approximately 11km*11km=121km^2) and clouds at any height above the surface are considered. The parameter is given in %.

Total cloud cover is available for the analysis and the forecast time steps. For the forecast, the value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

Please note that there is a problem with this parameter. Check section 2.3.1.2 Archiving issues for more details when using this parameter.

3.1.1.26 Snow density

Snow density is the snow mass per unit of volume. Hence, the parameter is given in kg/m^3. It is given as the mean for a grid point (approximately $11km*11km=121km^2$). Grid points without snow have missing values.

Snow density is available for the analysis and the forecast time steps up to forecast hour six. For the forecast, the value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.1.27 Snow depth water equivalent

Snow depth water equivalent expresses the snow depth in kg of snow over one square meter [kg/m^2]. The unit corresponds to 1 mm of water equivalent. It is given as the mean for a grid point (approximately $11km*11km=121km^2$).

Snow depth water equivalent is available for the analysis and the forecast time steps up to forecast hour six. For the forecast, the value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.1.28 Snow fall water equivalent

Snow fall water equivalent expresses the snow fall in kg of snow over one square meter [kg/m^2]. The unit corresponds to 1 mm of water equivalent. It is given as the mean for a grid point (approximately $11km*11km=121km^2$).

Snow fall water equivalent is only available for the forecast time steps. It is an accumulated parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 24h-forecast has the accumulated snow fall water equivalent over 24 hours.

3.1.1.29 Land-Sea mask

The land-sea mask is a field that contains, for every grid point (approximately $11km*11km=121km^2$), the proportion of land in the grid box. The parameter is dimensionless and the values are between 0 (sea) and 1 (land).

The land-sea mask is constant in time and the field is available for every analysis.

Note that the UERRA-HARMONIE land-sea mask does not contain all lakes. The lakes are treated separately by the model system.

3.1.1.30 Orography

The orography is the height of the terrain with respect to the model defined globe. Each grid point (approximately $11km*11km=121km^2$) has one value representing the mean over the grid point domain. The orography is given as geopotential height in meter [gpm].

The orography is constant in time and the field is available for every analysis.

3.1.1.31 Surface roughness

The surface roughness describes the aerodynamic roughness length (over land). Each grid point (approximately $11km*11km=121km^2$) has one value representing the mean over the grid point. The surface roughness is given in meter [m].

The surface roughness is available for the analysis and the forecast time steps up to forecast hour six. For the forecast, the value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.2 Parameters on height levels

Metadata	
Horizontal coverage	The model domain spans from northern Africa beyond the northern tip of Scandinavia. In the west it ranges far into the Atlantic ocean and in the east it reaches to the Ural. Herewith, it covers entire Europe. See Figure for an overview of the model domain.
Horizontal resolution	11x11 km^2.
Vertical coverage	11 height levels are available which are between 15-500m.
Vertical levels	15m, 30m, 50m, 75m, 100m, 150m, 200m, 250m, 300m, 400m, 500m
Temporal coverage	1961-01-01 00:00 – close to real time (monthly updates but with a delay to real time of about four months)
Temporal resolution	Analysis are available at 00, 06, 12, 18 UTC. The forecast length is depending on the cycle. Cycles initialized at 00 and 12 UTC have forecasts saved at 1, 2, 3, 4, 5, 6, 9, 12, 15, 18, 21, 24, 30. Cycles initialized at 06 and 18 UTC have forecasts saved at 1, 2, 3, 4, 5, 6. See 2.1.3 (Fig. 6) for more details.
Data type and format	Gridded data in GRIB2
Grid	Lambert conformal conic grid with 565x565 grid points

Table 2: Overview parameters on height levels

	Parameter	Unit	GRIB code	Analysis 0, 6, 12, 18	forecast 1,2,3,…
1.	Wind speed	m/s	10	yes	yes
2.	Wind direction	degrees	3031	yes	yes
3.	Pressure	Pa	54	yes	yes
4.	Specific cloud liquid water content	kg/kg	246	-	yes
5.	Specific cloud ice water content	kg/kg	247	-	yes
6.	Relative humidity	%	157	yes	yes
7.	Temperature	K	130	yes	yes

3.1.2.1 Wind speed

Wind speed is the wind speed valid for a grid point (approximately 11km*11km=121km^2) determined for a certain height (15m-500m) above the surface. The parameter is given in m/s. It is computed from both the zonal (u) and the meridional (v) wind components by

$$wind\ speed = \sqrt{u^2 + v^2}$$

The wind speed is available for the analysis and the forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.2.2 Wind direction

The wind direction is the wind direction valid for a grid point (approximately 11km*11km=121km^2) determined for a certain height (15m-500m) above the surface. The parameter is given in degrees ranging from 0-360. Here, 0° means a northerly wind and 90° indicates an easterly wind.
The wind direction is available for the analysis and the forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.2.3 Pressure

The pressure is the air pressure at a certain height (15m-500m) above the surface valid for a grid point (approximately 11km*11km=121km^2). The parameter is given in Pascal [Pa].
The pressure is available for the analysis and the forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.2.4 Specific cloud liquid water content

Specific cloud liquid water content is the grid-box mean (approximately 11km*11km=121km^2) liquid water content (mass of condensate / mass of moist air) on a height level. It is given in kg/kg.
The parameter is only available for forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.2.5 Specific cloud ice water content

Specific cloud ice water content is the grid-box mean (approximately 11km*11km=121km^2) ice water content (mass of condensate / mass of moist air) on a height level. It is given in kg/kg.
The parameter is only available for forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.2.6 Relative humidity

The relative humidity is the modelled humidity valid for a grid point (approximately 11km*11km=121km^2) determined at a certain height (15m-500m) above the surface. The parameter is given % ranging from 0-100. 0% means that the air is totally dry whereas 100% indicates that the air is saturated with water vapour. The saturation is defined with respect to saturation of the mixed phase, i.e. with respect to saturation over ice below -23°C and with respect to saturation over water above 0°C. In the regime in between a quadratic interpolation is applied.
Surface air relative humidity is available for the analysis and the forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.
Please check section 2.3 Model specific issues when using this parameter.

3.1.2.7 Temperature

The temperature is the air temperature valid for a grid point (approximately 11km*11km=121km^2) determined at a certain height (15m-500m) above the surface. The parameter is given in Kelvin [K].
The temperature is available for the analysis and the forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.3 Parameters on pressure levels

Metadata	
Horizontal coverage	The model domain spans from northern Africa beyond the northern tip of Scandinavia. In the west it ranges far into the Atlantic ocean and in the east it reaches to the Ural. Herewith, it covers entire Europe. See Figure for an overview of the model domain.
Horizontal resolution	11x11 km^2.
Vertical coverage	24 pressure levels are stored at levels 1000-10hPa.
Vertical levels [hPa]	1000, 975, 950, 925, 900, 875, 850, 825, 800, 750, 700, 600, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20, 10
Temporal coverage	1961-01-01 00:00 – close to real time (monthly updates but with a delay to real time of about four months)
Temporal resolution	Analysis are available at 00, 06, 12, 18 UTC. The forecast length is depending on the cycle. Cycles initialized at 00 and 12 UTC have forecasts saved at 1, 2, 3, 4, 5, 6, 9, 12, 15, 18, 21, 24, 30. Cycles initialized at 06 and 18 UTC have forecasts saved at 1, 2, 3, 4, 5, 6. See 2.1.3 (Fig. 6) or more details.
Data type and format	Gridded data in GRIB2
Grid	Lambert conformal conic grid with 565x565 grid points

Table 3: Overview parameters on pressure levels

	Parameter	Unit	GRIB code	Analysis 0, 6, 12, 18	forecast 1,2,3,…
1.	Cloud cover	%	260257	-	yes
2.	Specific cloud liquid water content	kg/kg	246	-	yes
3.	Specific cloud ice water content	kg/kg	247	-	yes
4.	Relative humidity	%	157	yes	yes
5.	Temperature	K	130	yes	yes
6.	U-component of wind	m/s	131	yes	yes
7.	V-component of wind	m/s	132	yes	yes
8.	Geopotential	m^2/s^2	129	yes	yes
9.	Geopotential height	gpm	156	yes	yes

3.1.3.1 Cloud cover

Cloud cover is the percentage of sky covert with clouds. It is valid for a grid point (approximately 11km*11km=121km^2) at the corresponding height. The parameter is given in %.
Total cloud cover is only available for the forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.3.2 Specific cloud liquid water content
The specific cloud liquid water content is the grid-box mean (approximately 11km*11km=121km^2) mass of condensate / mass of moist air on a pressure level. The parameter is given in kg/kg.
Specific cloud liquid water content is only available for the forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.
3.1.3.3 Specific cloud ice water content
The specific cloud ice water content is the grid-box mean (approximately 11km*11km=121km^2) mass of condensate / mass of moist air on a pressure level. The parameter is given in kg/kg.
Specific cloud ice water content is only available for the forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.
3.1.3.4 Relative humidity
The relative humidity is the modelled humidity valid for a grid point (approximately 11km*11km=121km^2) at the corresponding height. The parameter is given % ranging from 0-100. 0% means that the air is totally dry whereas 100% indicates that the air is saturated with water vapour. The saturation is defined with respect to saturation of the mixed phase, i.e. with respect to saturation over ice below -23°C and with respect to saturation over water above 0°C. In the regime in between a quadratic interpolation is applied.
Relative humidity is available for the analysis and the forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.
Please check section 2.3 Model specific issues when using this parameter.
3.1.3.5 Temperature
The temperature is the model temperature valid for a grid point (approximately 11km*11km=121km^2) at the corresponding height. The parameter is given in Kelvin [K].
Temperature is available for the analysis and the forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.
3.1.3.6 U-component of wind
The U-component of wind or U-velocity is the zonal component of the wind valid for a grid point (approximately 11km*11km=121km^2) at the corresponding height. The parameter is given in m/s. By model convention westerly wind (blowing from the west to the east) are positive.
U-velocity is available for the analysis and the forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.
3.1.3.7 V-component of wind
The V-component of wind or V-velocity is the meridional component of the wind valid for a grid point (approximately 11km*11km=121km^2) at the corresponding height. The parameter is given in m/s. By model convention southerly wind (blowing from the south to the north) are positive.
V-velocity is available for the analysis and the forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.
3.1.3.8 Geopotential
The geopotential is the potential energy of unit mass at this pressure level relative to the sea level. It is valid for a grid point (approximately 11km*11km=121km^2) and it is given in m^2/s^2.
The geopotential is available for the analysis and the forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.3.9 Geopotential height

The geopotential height is the altitude of the given pressure level in the atmosphere relative to the sea level. It is valid for a grid point (approximately 11km*11km=121km^2) and it is given in units proportional to the geopotential. Therefore, the unit is called geopotential meter [gpm].
The geopotential height is available for the analysis and the forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.
Note: The Geopotential and the geopotential height are related by a factor and are thus redundant.
Geopotential = Geopotential height / gravitational acceleration (g = 9.81 m/s^2)
During the UERRA production until 2015 both geopotential and geopotential heights were archived due to an earlier confusion of the two in the design document. In the production since 2016 only the intended one, geopotential, is maintained. However, the geopotential can be computed with the above given formula at any time.

3.1.4 Parameters on model levels

Metadata	
Horizontal coverage	The model domain spans from northern Africa beyond the northern tip of Scandinavia. In the west it ranges far into the Atlantic ocean and in the east it reaches to the Ural. Herewith, it covers entire Europe. See Figure for an overview of the model domain.
Horizontal resolution	11x11 km^2.
Vertical coverage	65 model levels from the surface to the model top at 10hPa.
Vertical levels	1, 2, 3, 4,… , 63, 64, 65
Temporal coverage	1961-01-01 00:00 – close to real time (monthly updates but with a delay to real time of about four months)
Temporal resolution	Analysis are available at 00, 06, 12, 18 UTC. Forecasts are not saved for the parameters on model levels.
Data type and format	Gridded data in GRIB2
Grid	Lambert conformal conic grid with 565x565 grid points

Table 4: Overview parameters on model levels

	Parameter	Unit	GRIB code	Analysis 0, 6, 12, 18	forecast 1,2,3,…
1.	Specific humidity	kg/kg	133	yes	-
2.	Temperature	K	130	yes	-
3.	U-velocity	m/s	131	yes	-
4.	V-velocity	m/s	132	yes	-

3.1.4.1 Specific humidity

The specific humidity is the mass of water vapor per unit mass of air valid for a grid point (approximately 11km*11km=121km^2) at the corresponding model level. The parameter is given in kg/kg.

Only analyses are stored for parameters on model levels. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.4.2 Temperature

The temperature is the model temperature valid for a grid point (approximately 11km*11km=121km^2) at the corresponding model level. The parameter is given in Kelvin [K].
Only analyses are stored for parameters on model levels. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.4.3 U-velocity

The U-velocity is the zonal component of the wind valid for a grid point (approximately 11km*11km=121km^2) at the corresponding model level. The parameter is given in m/s. By model convention westerly wind (blowing from the west to the east) are positive.
Only analyses are stored for parameters on model levels. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.4.4 V-velocity

The V-velocity is the meridional component of the wind valid for a grid point (approximately 11km*11km=121km^2) at the corresponding model level. The parameter is given in m/s. By model convention southerly wind (blowing from the south to the north) are positive.
Only analyses are stored for parameters on model levels. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.5 Soil parameters

Metadata	
Horizontal coverage	The model domain spans from northern Africa beyond the northern tip of Scandinavia. In the west it ranges far into the Atlantic ocean and in the east it reaches to the Ural. Herewith, it covers entire Europe. See Figure for an overview of the model domain.
Horizontal resolution	11x11 km^2.
Vertical coverage	3 levels of the soil model.
Vertical levels	1, 2, 3
Temporal coverage	1961-01-01 00:00 – close to real time (monthly updates but with a delay to real time of about four months)
Temporal resolution	Analysis are available at 00, 06, 12, 18 UTC. Forecasts are saved hourly up to six hours.
Data type and format	Gridded data in GRIB2
Grid	Lambert conformal conic grid with 565x565 grid points

Table 5: Overview soil parameters

	Parameter	Unit	GRIB code	Analysis 0, 6, 12, 18	forecast 1,2,3,4,5,6

1.	Volumetric soil moisture	m^3/m^3	260199	yes	yes
2.	Soil temperature	K	260360	yes	yes

3.1.5.1 Volumetric soil moisture
The volumetric soil moisture is the amount of water in a cubic meter soil valid for a grid point (approximately 11km*11km=121km^2) in the corresponding soil level. The parameter is given in m^3/m^3.
The parameter is available for analysis and forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.
3.1.5.2 Soil temperature
The soil temperature is the model temperature valid for a grid point (approximately 11km*11km=121km^2) at the corresponding soil level. The parameter is given in Kelvin [K].
The parameter is available for analysis and forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.1.6 The UERRA-HARMONIE grid description

Below are the essential parameters describing the grid and the used Lambert Conformal Conic projection. More information about the grid and coordinates can be found in the FAQ in section 2.4.

Number of points along x-axis: 565
Number of points along y-axis: 565
X-direction grid length: 11000 m
Y-direction grid length: 11000 m

Projection: Lambert Conformal Conic
Central meridian: 8
Standard parallel 1: 48
Standard parallel 2: 48
Latitude of origin: 48

Earth assumed spherical with radius: 6371229 m

Latitude of first grid point in degrees: 17.612
Longitude of first grid point in degrees: 341.68

3.2 The MESCAN-SURFEX system

3.2.1 Surface parameters

Metadata	
Horizontal coverage	The model domain spans from northern Africa beyond the northern tip of Scandinavia. In the west it ranges far into the Atlantic ocean and in the east it reaches to the Ural. Herewith, it covers entire Europe. See Figure for an overview of the model domain.
Horizontal resolution	5.5x5.5 km^2.
Vertical coverage	Each surface parameter is valid for one vertical level and three different (near) surface levels exists: - surface: atmospheric boundary with the ground or water surface - 2m: 2m above the surface - 10m: 10m above the surface
Vertical levels	One level only
Temporal coverage	1961-01-01 00:00 – close to real time (monthly updates but with a delay to real time of about four months)
Temporal resolution	Analysis are available at 00, 06, 12, 18 UTC. Forecasts are computed up to 6 hours and the output frequency is hourly though some of the parameters are saved only for the six hour forecasts.
Data type and format	Gridded data in GRIB2
Grid	Lambert conformal conic grid with 1069x1069 grid points

Table 6: Overview surface parameters from the MESCAN-SURFEX system

	Parameter	Unit	GRIB code	Analysis 0, 6, 12, 18	Forecast 1,2,3,4,5,6	Height
1.	2m relative humidity	%	260242	yes	only at 6h	2m
2.	Total precipitation	kg/m^2	228228	only available at 6UTC (24h accumulate d)	only at 6h	surface
3.	10m wind speed	m/s	207	yes	only at 6h	10m
4.	10m wind direction	degrees	260260	yes	only at 6h	10m
5.	2m temperature	K	167	yes	only at 6h	2m
6.	Skin temperature	K	235	-	yes	surface
7.	Albedo	0-1	260509	-	yes	surface
8.	Time-integrated surface latent heat flux	J/m^2	147	-	yes	surface
9.	Time-integrated surface sensible heat flux	J/m^2	146	-	yes	surface

10.	Time-integrated surface direct solar radiation	J/m^2	260264	-	only at 6h	surface
11.	Time-integrated surface net solar radiation	J/m^2	176	-	yes	surface
12.	Time-integrated surface solar radiation downwards	J/m^2	169	-	yes	surface
13.	Time-integrated surface net thermal radiation	J/m^2	177	-	yes	surface
14.	Time-integrated surface thermal radiation downwards	J/m^2	175	-	yes	surface
15.	Surface pressure	Pa	134	-	only at 6h	surface
16.	Surface runoff	kg/m^2	174008	-	yes	surface
17.	Soil heat flux	W/m^2	260364	-	yes	surface
18.	Surface roughness	m	173	-	yes	surface
19	Snow depth	m	3066	-	yes	surface
20.	Snow density	kg/m^3	33	-	yes	surface
21.	Snow depth water equivalent	kg/m^2	228141	-	yes	surface
22.	Snow fall water equivalent	kg/m^2	228144	-	yes	surface
23	Percolation	kg/m^2	260430	-	yes	surface
24.	Land-sea mask	dimension-less	172	yes	-	surface
25.	Orography	gpm	228002	yes	-	surface

3.2.1.1 2m relative humidity

The 2m relative humidity (parameter name surface air relative humidity) is the modelled humidity valid for a grid point (approximately 5.5km*5.5km) determined for a height of 2m above the surface. The parameter is the relation between actual humidity and saturation humidity given in %. Values are in the interval [0,100]. 0% means that the air is totally dry whereas 100% indicates that the air is saturated with water vapour. The saturation is defined with respect to saturation of the mixed phase, i.e. with respect to saturation over ice below -23°C and with respect to saturation over water above 0°C. In the regime in between a quadratic interpolation is applied.
Surface air relative humidity is available for the analysis and the 6h-forecast. For the forecast, the value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.2.1.2 Total precipitation

Total precipitation is the amount of precipitation falling onto the ground/water surface. It includes all kind of precipitation forms as convective precipitation, large scale precipitation, liquid and solid precipitation. The amount is valid for a grid box and has the unit kg/m^2.
The total precipitation is available for the analysis at 06UTC with accumulated values over 24 hours. For the forecast time steps, total precipitation is available only for the six hour forecast. It is an accumulated parameter meaning that it is accumulated from the beginning of the forecast – hence, accumulated over 6 hours.

Unfortunately, there is a quality problem for this parameter. Please check section 2.3.2.2 when using this parameter.

3.2.1.3 10m wind speed

The 10m wind speed is the wind speed valid for a grid point (approximately 5.5km*5.5km) determined for a height of 10m above the surface. The parameter is given in m/s. It is computed from both the zonal (u) and the meridional (v) wind components by

$$wind\ speed = \sqrt{u^2 + v^2}$$

The 10 metre wind speed is available for the analysis and the forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.2.1.4 10m wind direction

The 10 metre wind direction is the wind direction valid for a grid point (approximately 5.5km*5.5km) determined for a height of 10m above the surface. The parameter is given in degrees ranging from 0-360. Here, 0° means a northerly wind and 90° indicates an easterly wind.
The 10 metre wind direction is available for the analysis and the forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.2.1.5 2m temperature

The 2m temperature with the parameter name surface air temperature is the model temperature valid for a grid point (approximately 5.5km*5.5km) determined for a height of 2m above the surface. The parameter is given in Kelvin [K].
Surface air temperature is available for the analysis and the 6h-forecast. For the forecast, the value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.
Unfortunately, there is a quality problem for this parameter. Please check section 2.3.2.1 when using this parameter.

3.2.1.6 Skin temperature

The skin temperature is the model temperature valid for a grid point (approximately 5.5km*5.5km) determined for the boundary surface to the atmosphere, both ground and water surfaces. The parameter is given in Kelvin [K].
Skin temperature is available for the forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.2.1.7 Albedo

The albedo is the amount of radiation which is reflected at the valid for a grid point (approximately 5.5km*5.5km). It is determined for the surface to the atmosphere, both for ground and water surfaces. The parameter is given in %. Small values mean that large amounts of the radiation are absorbed whereas large values mean that more radiation is reflected.
Albedo is available for the analysis and the forecast time steps up to forecast hour six. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.
Albedo is defined only over land and has missing values over sea and lakes.

3.2.1.8 Time-integrated surface latent heat flux

The surface latent heat flux is the exchange of latent heat (due to phase transitions: evaporation, condensation) with the surface (ground and water) through turbulent diffusion. It is given as a mean

for a grid point (approximately 5.5km*5.5km). The parameter is given in J/m^2. By model convention downward fluxes are positive.
Surface latent heat flux is only available for forecast time steps up to forecast hour six. It is an accumulated (time-integrated) parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 6h-forecast has the accumulated heat flux over 6 hours.

3.2.1.9 Time-integrated surface sensible heat flux

The surface sensible heat flux is the exchange of heat (no phase transition) with the surface (ground and water) through turbulent diffusion. It is given as a mean for a grid point (approximately 5.5km*5.5km). The parameter is given in J/m^2. By model convention downward fluxes are positive.
Surface sensible heat flux is only available for forecast time steps up to forecast hour six. It is an accumulated (time-integrated) parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 6h-forecast has the accumulated heat flux over 6 hours.

3.2.1.10 Time-integrated surface direct solar radiation

The surface direct solar radiation is the amount of direct solar (short-wave) radiation reaching the surface (ground and water). It is given as a mean for a grid point (approximately 5.5km*5.5km). The parameter is given in J/m^2. By model convention downward fluxes are positive.
Surface direct solar radiation is only available for forecast time steps. It is an accumulated (time-integrated) parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 6h-forecast has the accumulated radiation over 6 hours.

3.2.1.11 Time-integrated surface net solar radiation

The surface net solar radiation is the amount of solar (short-wave) radiation that is absorbed at the surface (ground and water). It is computed

*Surface net solar radiation = surface solar radiation downwards * (1 – albedo)*

It is given as a mean for a grid point (approximately 5.5km*5.5km). The parameter is given in J/m^2. By model convention downward fluxes are positive.
Surface net solar radiation is only available for forecast time steps. It is an accumulated (time-integrated) parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 6h-forecast has the accumulated radiation over 6 hours.

3.2.1.12 Time-integrated surface solar radiation downwards

The surface solar radiation downward is the amount of solar (short-wave) radiation reaching the surface (ground and water). It is given as a mean for a grid point (approximately 5.5km*5.5km). The parameter is given in J/m^2. By model convention downward fluxes are positive.
Surface solar radiation downwards is only available for forecast time steps. It is an accumulated (time-integrated) parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 6h-forecast has the accumulated radiation over 6 hours.
Please check section 2.3 Model specific issues when using this parameter.

3.2.1.13 Time-integrated surface net thermal radiation

The surface net thermal radiation is the amount of thermal (long-wave) radiation that is absorbed at the surface (ground and water). It is computed

*Surface net thermal radiation = surface thermal radiation downwards * (1 – albedo)*

It is given as a mean for a grid point (approximately 5.5km*5.5km). The parameter is given in J/m^2. By model convention downward fluxes are positive.
Surface net thermal radiation is only available for forecast time steps. It is an accumulated (time-integrated) parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 6h-forecast has the accumulated radiation over 6 hours.

3.2.1.14 Time-integrated surface thermal radiation downwards

The surface thermal radiation downward is the amount of thermal (long-wave) radiation reaching the surface (ground and water). It is given as a mean for a grid point (approximately 5.5km*5.5km). The parameter is given in J/m^2. By model convention downward fluxes are positive.

Surface thermal radiation downwards is only available for forecast time steps. It is an accumulated (time-integrated) parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 6h-forecast has the accumulated radiation over 6 hours.

3.2.1.15 Surface pressure

The surface pressure is the air pressure at the surface (ground and water) valid for a grid point (approximately 5.5km*5.5km). The parameter is given in Pascal [Pa].

Mean sea level pressure is available for the analysis and the forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.2.1.16 Surface runoff

Surface runoff is the lateral water flow occurring at the surface of a grid point (approximately 5.5km*5.5km). The parameter is given in kg/m^2.

Surface runoff is only available for forecast time steps. It is an accumulated (time-integrated) parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 6h-forecast has the accumulated surface runoff over 6 hours.

3.2.1.17 Soil heat flux

The soil heat flux is the energy received by the soil to heat it per unit of surface and time. It is valid for a grid point and has the unit W/m^2. The Soil heat flux is positive when the soil receives energy (warms) and negative when the soil loses energy (cools).

Soil heat flux is only available for forecast time steps. It is an accumulated (time-integrated) parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 6h-forecast has the accumulated surface runoff over 6 hours.

3.2.1.18 Surface roughness

The surface roughness describes the aerodynamic roughness length (over land). Each grid point (approximately 5.5km*5.5km) has one value representing the mean over the grid point. The surface roughness is given in meter [m].

The surface roughness is available for the analysis and the forecast time steps up to forecast hour six. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step. The surface roughness is depending on the seasonal variations of the vegetation as well as snow cover.

3.2.1.19 Snow depth

Snow depth is the average snow height for a grid point (approximately 5.5km*5.5km). Snow depth is given in meter [m].

Snow depth is available for the forecast time steps up to forecast hour six. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.2.1.20 Snow density
Snow density is the snow mass per unit of volume. Hence, the parameter is given in kg/m^3. It is given as the mean for a grid point (approximately 5.5km*5.5km). Grid points without snow have missing values.
Snow density is available for the forecast time steps up to forecast hour six. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.
3.2.1.21 Snow depth water equivalent
Snow depth water equivalent expresses the snow depth in kg of snow over one square meter [kg/m^2]. The unit corresponds to 1 mm of water equivalent. It is given as the mean for a grid point (approximately 5.5km*5.5km).
Snow depth water equivalent is available for the analysis and the forecast time steps up to forecast hour six. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.
3.2.1.22 Snow fall water equivalent
Snow fall water equivalent expresses the snow fall in kg of snow over one square meter [kg/m^2]. The unit corresponds to 1 mm of water equivalent. It is given as the mean for a grid point (approximately 5.5km*5.5km).
Snow fall water equivalent is only available for the forecast time steps. It is an accumulated parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 6h-forecast has the accumulated snow fall water equivalent over 6 hours.
3.2.1.23 Percolation
Percolation is the soil total column lateral water flow and bottom soil drainage. It is given as the mean for a grid point (approximately 5.5km*5.5km) and has the unit kg/m^2.
Percolation is only available for the forecast time steps. It is an accumulated parameter meaning that it is accumulated from the beginning of the forecast. For instance, the 6h-forecast has the accumulated water flow over 6 hours.
3.2.1.24 Land-Sea mask
The land-sea mask is a field that indicates, for every grid point (approximately 5.5km*5.5km), whether the grid box is a land point or a wet point. The parameter is dimensionless and the values are 0 (for wet points) and 1 (for land points).
Please check section 2.3 Model specific issues when using this parameter.
3.2.1.25 Orography
The orography is the height of the terrain with respect to the model defined globe. Each grid point (approximately 5.5km*5.5km) has one value representing the mean over the grid point domain. The orography is given as geopotential height in meter [gpm].
The orography is constant in time and the field is available for every analysis.

3.2.2 Soil parameters

Metadata	
Horizontal coverage	The model domain spans from northern Africa beyond the northern tip of Scandinavia. In the west it ranges far into the Atlantic ocean and in the east it reaches to the Ural. Herewith, it covers entire Europe.

	See Figure for an overview of the model domain.
Horizontal resolution	5.5x5.5 km^2.
Vertical coverage	14 levels of the soil model from the surface to a depth of 12m.
Vertical levels as depth of the considered layer in meter [m]	0.01, 0.04, 0.1, 0.2, 0.4, 0.6, 0.8, 1, 1.5, 2, 3, 5, 8, 12
Temporal coverage	1961-01-01 00:00 – close to real time (monthly updates but with a delay to real time of about four months)
Temporal resolution	Analysis are available at 00, 06, 12, 18 UTC.
Data type and format	Gridded data in GRIB2
Grid	Lambert conformal conic grid with 1069x1069 grid points

Table 5: Overview soil parameters from the MESCAN-SURFEX system

	Parameter	Unit	GRIB code	Analysis 0, 6, 12, 18	forecast 1,2,3,…
1.	Liquid non-frozen volumetric soil water	m^3/m^3	260210	-	yes
2.	Volumetric soil moisture	m^3/m^3	260199	-	yes
3.	Soil temperature	K	260360	-	yes
4.	Volumetric wilting point	m^3/m^3	260200	yes	-
5.	Volumetric transpiration stress-onset	m^3/m^3	260211	yes	-

3.2.2.1 Liquid non-frozen volumetric soil water

The liquid non-frozen volumetric soil water is the amount of liquid water in a cubic meter soil valid for a grid point (approximately 5.5km*5.5km) in the corresponding soil level. The parameter is given in m^3/m^3.

The parameter is only available for forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.2.2.2 Volumetric soil moisture

The volumetric soil moisture is the amount of water in a cubic meter soil valid for a grid point (approximately 5.5km*5.5km) in the corresponding soil level. The parameter is given in m^3/m^3.

The parameter is only available for forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

To interpret soil water and to compare different models the Soil Wetness Index (SWI) is used:

SWI = (soil_water - wilting_point) / (field_capacity - wilting_point).

3.2.2.3 Soil temperature

The soil temperature is the model temperature valid for a grid point (approximately 5.5km*5.5km) at the corresponding soil level. The parameter is given in Kelvin [K].

The parameter is only available for forecast time steps. The value is instantaneous meaning that it is valid for the last time step of the integration at the issued time step.

3.2.2.4 Volumetric wilting point
The volumetric wilting point corresponds to the soil water content at which plants wilt and can no longer recover. It is given for a grid point (approximately 5.5km*5.5km) in the corresponding soil level. The parameter is given in m^3/m^3.
The parameter is only available for the analyses.

3.2.2.5 Volumetric transpiration stress-onset (soil moisture)
The volumetric transpiration stress-onset (or volumetric field capacity) corresponds to the soil water content after the soil has been saturated and allowed to drain freely. It is given for a grid point (approximately 5.5km*5.5km) in the corresponding soil level. The parameter is given in m3/m3.
The parameter is only available for the analyses.

3.2.3 The MESCAN-SURFEX grid description
Below are the essential parameters describing the grid and the used Lambert Conformal Conic projection. More information about the grid and coordinates can be found in the FAQ in section 2.4.

Number of points along x-axis: 1069
Number of points along y-axis: 1069
X-direction grid length: 5500 m
Y-direction grid length: 5500 m

Projection: Lambert Conformal Conic
Central meridian: 8
Standard parallel 1: 50
Standard parallel 2: 50
Latitude of origin: 50

Earth assumed spherical with radius: 6371229 m

Latitude of first grid point in degrees: 20.292
Longitude of first grid point in degrees: 342.514

4. Data access

The UERRA reanalysis datasets (in total more than 700 Tb) are freely available. The analyses are available from the Copernicus Climate Data Store (CDS) at https://cds.climate.copernicus.eu/. Here, users have even the possibility to download the data as netCDF as well as using the data in the coupled toolbox. Since autumn 2019, it is possible to access all fields (including forecast fields as well as the parameters on model levels) via the Climate Data Store Application Program Interface (CDS API).
In addition, user might be interested to access the data via the MARS archive of ECMWF at https://apps.ecmwf.int/datasets/data/uerra.
Details for the methods are given below.

4.1 Data access via CDS

After registration at the CDS there are several options to download and visualize the data. An introduction to the CDS and the CDS Toolbox is available as a lesson in the Copernicus User Learning Services at https://uls.climate.copernicus.eu/browse-lessons?packageId=1148 .

The options for data access are:

1) Download the data via a form in your web browser.
2) Use CDS Toolbox to write and execute an extraction script.
3) Install the python package cdsapi and get the data directly onto your computer.

Each of these options is further described below. Please note that at the time of writing only the analysis data are accessible via the CDS and Toolbox (option 1 and 2 above). To access all the data, including for example the hourly forecasts, you need to access the data via CDS API (option 3 above), see section 4.2 for more information about this.

Option 1 is very straight forward; just choose one of the UERRA datasets available in the CDS from: https://cds.climate.copernicus.eu/cdsapp#!/search?type=dataset&text=uerra
then click the "Download data" tab in the browser, fill in and submit the form. Below the form are also buttons for showing the corresponding Toolbox request (see option 2 below) and API request (see option 3 below). The result of these requests will be files in GRIB or NetCDF format – depending on your choice.

Option 2 is mainly aimed at further processing of the data in the CDS Toolbox. This involves writing some code in the CDS Toolbox editor. You can take this code from the Toolbox request generated in option 1 above. As mentioned, this method is mainly for using data in the CDS Toolbox, but you could also use it to download data. There is an example retrieval script at (logging in required) https://cds.climate.copernicus.eu/toolbox-editor/examples/01-retrieve-data.

Option 3 is the preferred method for downloading larger amounts of data. To use this method you need to install the python package cdsapi on your computer. This is described at https://cds.climate.copernicus.eu/api-how-to.

After installing the CDS API you can execute python scripts to retrieve UERRA data. In section 4.1.1 is an example for retrieving the UERRA-HARMONIE 2m temperature analyses. The request is based

on code generated by the CDS web form (see option 1 above). In section 4.1.2 is an example for retrieving the hourly UERRA-HARMONIE 2m temperature forecast. Note that the complete data, such as forecasts, are only available via CDS API and not through the CDS web interface. The request syntax is also slightly different for the complete data. It is based on the ECMWF MARS retrieval system syntax.
Examples of requests using the MARS syntax can be generated via the web forms available at https://apps.ecmwf.int/datasets/data/uerra/
The full set of variables that are available can be found at https://software.ecmwf.int/wiki/display/UER/Parameters

These scripts, and some other examples, are also available for download via the open GitHub at https://git.smhi.se/C3S_322_Lot1/C3S_322_Lot1_user_examples.

4.1.1 CDS API analysis data example, UERRA-HARMONIE 2m-temperature analysis

```
#!/usr/bin/env python

import cdsapi
import calendar

c = cdsapi.Client()

def retrieve_uerra():
  """
    A function to demonstrate how to iterate over several years and months etc
    for a particular UERRA request for origin UERRA-HARMONIE.
    Change the variables below to adapt the iteration to your needs.
    You can use the variable 'targetFile' to organise the requested data in files as you wish.
    In the example below the data are organised in files per month.
  """
  yearStart = 2015
  yearEnd   = 2015
  monthStart =  1
  monthEnd   = 12
  for year in list(range(yearStart, yearEnd + 1)):
    for month in list(range(monthStart, monthEnd + 1)):
      numberOfDays = calendar.monthrange(year, month)[1]
      targetFile = "ofile_%04d%02d.grb" % (year, month)
      requestDates = ['{:02}'.format(i) for i in range(1, numberOfDays+1)]
      requestMonth = '{:02}'.format(month)
      requestYear = '{:04}'.format(year)
      uerra_request(requestYear, requestMonth, requestDates, targetFile)
```

```
def uerra_request(reqYear, reqMonth, reqDates, target):
    """
        A UERRA request for 2 metre temperature every 6th hour.
        Origin uerra_harmonie, surface level, analysis fields.
        Request cost per day is 4 fields, 1.8 Mbytes.
    """

    c.retrieve(
        'reanalysis-uerra-europe-single-levels',
        {
            'format':'grib',
            'variable':'2m_temperature',
            'year':reqYear,
            'month':reqMonth,
            'day':reqDates,
            'time':['00:00','06:00','12:00','18:00'],
            'origin':'uerra_harmonie'
        },
        target)

if __name__ == '__main__':
    retrieve_uerra()
```

4.1.2 CDS API complete data example. UERRA-HARMONIE hourly 2m-temperature forecast

```
#!/usr/bin/env python

import cdsapi
import calendar

c = cdsapi.Client()

def retrieve_uerra():
    """
       A function to demonstrate how to iterate efficiently over several years and months etc
       for a particular UERRA request for origin SMHI.
       Change the variables below to adapt the iteration to your needs.
       You can use the variable 'targetFile' to organise the requested data in files as you wish.
       In the example below the data are organised in files per month.
    """
```

```
    yearStart = 2016
    yearEnd   = 2016
    monthStart =  1
    monthEnd   = 12
    for year in list(range(yearStart, yearEnd + 1)):
        for month in list(range(monthStart, monthEnd + 1)):
            startDate = '%04d%02d%02d' % (year, month, 1)
            numberOfDays = calendar.monthrange(year, month)[1]
            requestDates = ['{:04}'.format(year)+'{:02}'.format(month)+'{:02}'.format(i) for i in range(1,
numberOfDays+1)]
            targetFile = "ofile_%04d%02d.grb" % (year, month)
            uerra_request(requestDates, targetFile)


def uerra_request(requestDates, target):
    """
        A UERRA request for 2 metre temperature every hour.
        Origin SMHI, surface level, forecast fields.
        Request cost per day is 24 fields, 11 Mbytes.
    """
    c.retrieve(
        'reanalysis-uerra-europe-complete',
        {
         'class':'ur',
         'database':'external',
         'stream':'oper',
         'format':'grib',
         'type':'fc',
         'step':'1/2/3/4/5/6',
         'origin':'eswi',
         'date': requestDates,
         'expver':'prod',
         'levtype':'sfc',
         'param':'167',
         'time':'00/06/12/18'
         },
        target)


if __name__ == '__main__':
    retrieve_uerra()
```

4.2 Data access via MARS

The data are also available at the MARS archive of ECMWF at https://apps.ecmwf.int/datasets/data/uerra. This might be of interest for users who are familiar with the MARS system. However, in general, it is recommended to access the data via CDS or CDS API.

There are several options to download the data from ECMWFs archive. In any case, users will have to register themselves. There are two ways to register. First, users can register via a webpage and become a common user. Otherwise, users might register via their institution/country (as far as their institution or country is a member to ECMWF, respectively) and become a user with sophisticated rights.

The options for data access are:

1) You log into the ECMWF website and download the data manually through your web browser.
2) You log into ECMWFs super computer (you will need an institutional account) and extract the data directly from the archive.
3) You install the python package ecmwfapi on your machine and get the data directly onto your computer.

We consider version 3) as the most convenient method for users and focus therefore on it. In this document, we give one code example for potential data retrieval. The same script as well as other examples is available for download in the open GitHub at
https://git.smhi.se/C3S_322_Lot1/C3S_322_Lot1_user_examples
For instance, scripts to download the entire UERRA data set are provided.

The following request syntax is based on the ECMWF MARS retrieval system. These requests can for example be generated via web forms available at https://apps.ecmwf.int/datasets/data/uerra/

4.2.1 MARS example request

```
#!/usr/bin/env python

import calendar
from ecmwfapi import ECMWFDataServer
server = ECMWFDataServer()

def retrieve_uerra_eswi():
   """
    A function to demonstrate how to iterate efficiently over several years and months etc
    for a particular UERRA request for origin SMHI.
    Change the variables below to adapt the iteration to your needs.
    You can use the variable 'target' to organise the requested data in files as you wish.
    In the example below the data are organised in files per month.
```

```
    """
    yearStart = 2016
    yearEnd   = 2016
    monthStart =  1
    monthEnd   = 12
    for year in list(range(yearStart, yearEnd + 1)):
        for month in list(range(monthStart, monthEnd + 1)):
            startDate = '%04d%02d%02d' % (year, month, 1)
            numberOfDays = calendar.monthrange(year, month)[1]
            lastDate = '%04d%02d%02d' % (year, month, numberOfDays)
            target = "ofile_%04d%02d.grb" % (year, month)
            requestDates = (startDate + "/TO/" + lastDate)
            uerra_eswi_request(requestDates, target)

def uerra_eswi_request(requestDates, target):
    """
        A UERRA request for 2 metre temperature every hour.
        Origin SMHI, surface level, forecast fields.
        Request cost per day is 24 fields, 11 Mbytes.
    """
    server.retrieve({
        "class": "ur",
        "stream": "oper",
        "type": "fc",
        "dataset": "uerra",
        "origin" : "eswi",
        "date": requestDates,
        "expver": "prod",
        "levtype": "sfc",
        "param": "167",
        "target": target,
        "time": "00/06/12/18",
        "step": "1/2/3/4/5/6",
    })

if __name__ == '__main__':
    retrieve_uerra_eswi()
```

5. References/Further reading

5.1 References

- Bazile, E; R. Abida , A. Verelle, P. Le Moigne, C. Szczypta (2017): MESCAN-SURFEX surface analysis, deliverable D2.8 of the UERRA project, http://www.uerra.eu/publications/deliverable-reports.html
- Niermann, D. et al. (2017): Scientific report on assessment of regional analysis against independent data sets, deliverable D3.6 of the UERRA project, http://www.uerra.eu/publications/deliverable-reports.html
- Ridal, M.; S. Schimanke; S. Hopsch (2018): Documentation of the RRA system: UERRA, deliverable D322_Lot1.1.1.2 in the scope of the Copernicus service C3S_322_Lot1, available via Copernicus
- Soci, C., Bazile, E., Besson, F., & Landelius, T. (2016). High-resolution precipitation re-analysis system for climatological purposes. *Tellus A*, *68*
- Verver Gé (2017): User Guidance, deliverable D8.4 of the UERRA project, http://www.uerra.eu/publications/deliverable-reports.html

5.2 Further reading

Within the UERRA project, many reports were written describing the system and the data. It is recommend to have a look at these documents if you are looking for something specific. The collection of UERRA deliverables can be found here: http://www.uerra.eu/publications/deliverable-reports.html

For users, the reports listed below might be of specific interest:

- Deliverable 2.14: Reanalysis uncertainty evaluation
- Deliverable 3.8: User friendly synthesis report on evaluation and uncertainty of regional reanalyses
- Deliverable 4.5: Indices based on reanalysis data, including uncertainty information
- Deliverable 7.2: Training material on the use of reanalysis in climate services

You will find an overview of current atmospheric reanalyses activities at https://reanalyses.org/index.php/atmosphere

Copernicus Climate Change Service

ECMWF - Shinfield Park, Reading RG2 9AX, UK

Contact: info@copernicus-climate.eu