A high-resolution regional reanalysis for Europe. Part 2: 2D analysis of surface temperature, precipitation and wind

T. Landelius,* P. Dahlgren, S. Gollvik, A. Jansson and E. Olsson
Swedish Meteorological and Hydrological Institute, Norrköping, Sweden

*Correspondence to: T. Landelius, SMHI, Folkborgsvägen 17, 601 74 Norrköping, Sweden.
E-mail: tomas.landelius@smhi.se

The set-up and performance of the regional reanalysis for Europe with the HIgh-Resolution Limited-Area Model (HIRLAM) to a 3D grid-mesh with 22 km resolution for the years 1989–2010 have been presented in Part 1. This part describes how the 3D dataset is further downscaled and used as input for an analysis of a number of surface-related parameters: 2 m temperature, minimum and maximum daily temperatures, 10 m wind, and daily precipitation. The analysis is done on a 2D grid-mesh with 5 km grid spacing using the MESoscale ANalysis system (MESAN) for temperature and precipitation and a dynamical adaptation method (DYNAD) for the 10 m wind.

Results from MESAN and DYNAD are compared with observations and the HIRLAM 3D-Var reanalysis. A couple of cases with severe weather are studied to illustrate how such events are represented in the analyses. The comparisons show statistically significant added value in comparison to the HIRLAM reanalysis.

Key Words: reanalysis; data assimilation; downscaling; meteorological observations; regional climate

Received 19 May 2015; Revised 30 March 2016; Accepted 3 April 2016; Published online in Wiley Online Library 31 May 2016

1. Introduction

The purpose of the present articles is to describe the regional reanalysis for Europe that was done as part of the EU FP7 project ‘European Reanalysis and Observations for Monitoring’ (EURO4M). So far this is the only multidecadal regional reanalysis covering all of Europe. The goal of EURO4M was to develop the capacity for, and deliver, the best possible and most complete gridded climate time series and monitoring services covering all of Europe. Ongoing efforts within the follow-on project ‘Uncertainties in Ensembles of Regional Re-Analyses’ (UERRA) will result in an ensemble of European reanalyses at various resolutions, some of which will date back to 1960.

In Part 1 of this article the HIRLAM EURO4M reanalysis was described and evaluated. Adding value to a 3D reanalysis by supplementing it with a dedicated land surface reanalysis is becoming practice within reanalysis efforts worldwide (Dee et al., 2014). In order to further enhance and add detail to the EURO4M HIRLAM 3D-Var analysis, a 2D downsampling and analysis was performed.

In EURO4M the MESAN analysis system (Hägemark et al., 2000) was used to analyse a selection of surface parameters: daily precipitation (rr24), 2 m temperature (t2m), and its daily minimum and maximum (tn and tx respectively). MESAN is a system for operational mesoscale univariate analysis of selected meteorological parameters. The operational MESAN analyses are used for nowcasting and provide gridded input data to other models, e.g. concerning hydrology, air quality, atmospheric deposition, solar radiation and forest fire risks.

The 10 m winds (u10) were not analysed using MESAN. Instead the wind fields from HIRLAM were adjusted to the higher-resolution orography of the 2D grid using the method of dynamic adaptation (Zagar and Rakovec, 1999) with a simplified version of HIRLAM called DYNAD.

There are two factors contributing to the added value from the 2D reanalysis. First, adjustment of the surface parameters to a higher-resolution topography. The HIRLAM 3D fields are downscaled to the finer 2D grid in order to produce a first guess for the t2m analysis and an initial state for the u10 adaptation. This procedure is shown to result in a better t2m analysis than one based on a simple correction assuming a constant lapse rate. Second, more observations are used than in the 3D analysis. Additional databases were searched for t2m observations not present in the ECMWF MARS archive used by the HIRLAM 3D analysis. Neither precipitation nor minimum/maximum temperatures are assimilated in HIRLAM. Since the first guess to the rr24 analysis is a horizontally interpolated HIRLAM forecast, the added value of the 2D analysis of daily precipitation is all due to the introduction of unused observations. For tn and tx, the added value comes from the combined effect of incorporating unused observations and the adjustment to a higher-resolved orography.

In this article the reanalysis models, data assimilation methods, and observations used to produce the EURO4M 2D reanalysis with MESAN and DYNAD are described. The EURO4M 3D
and 2D reanalyses have now been extended and both cover the years 1979–2013. However, this happened at a time when the material for this article was almost completed. Hence, the resulting 2D reanalyses are here compared with observations and the EURO4M HIRLAM 3D-Var reanalysis for the years 1989–2010. Nevertheless, both the 3D and the 2D reanalyses can now be used to calculate climate normals for the 1981–2010 period.

Section 2 of this article describes the further downscaling and analysis of the 3D reanalysis using MESAN and DYNAD. The performance of the 2D reanalysis is evaluated in section 3 and its representation of extreme weather events is illustrated in section 4. The article ends with a summary and some outlooks regarding future plans for the dataset in section 5.

2. Downscaling and 2D analysis

The 2D reanalysis is performed at approximately 5 km compared to the 3D analysis which is carried out at about 20 km. For temperature and precipitation, the 2D analysis combines observations with a first guess in terms of a downscaled background field from the 3D regional system. The 10 m wind is dynamically adapted to the 5 km orography using DYNAD and a downscaled 3D model state for initialization. Since the MESAN analysis is univariate and not aimed at producing a balanced model state for a stable integration of an NWP forecast, it can use a more detailed orography and analyse entities that are otherwise hard to utilize in data assimilation for NWP. MESAN also employs regional variations of two factors such as land/sea mask and orography for its analysis.

2.1. Geometry

The horizontal grid in MESAN and DYNAD covers the same geographical region as HIRLAM and is expressed in an identical rotated latitude and longitude geometry. However, the resolution of the mesh is four times higher, resulting in a grid with 1286 \times 1361 points and 0.05° resolution (5.5 km). The maps presented in this article focus on the official EURO4M European area covering 25°W–45°E and 30–75°N.

2.2. Downscaling

The downscaling of the HIRLAM background fields is done differently for different parameters. The temperature fields are interpolated both horizontally and vertically taking into consideration differences in land/sea fractions between the two grids as well as the vertical profile of the modelled 3D atmosphere. This is also the case for the fields used as input to the dynamic adaptation DYNAD. On the other hand, downscaling of daily precipitation is done only by horizontal interpolation of the HIRLAM fields. Below the methods are described in more detail.

2.2.1. Temperature

The temperature is analysed eight times a day. In order to get the temperature forecast closest in time to the analysis hour (0000, 0300, 0600, 0900, 1200, 1500, 1800, and 2100 UTC) 3 h and 6 h forecasts from HIRLAM are used for the corresponding backgrounds: fc1800+0600, fc0900+0300, fc0000+0600, fc0600+0300, fc0000+0600, fc1200+0300, fc1200+0600, and fc1800+0300 UTC. Here ‘fcHH+LL’ denotes a forecast of length 3 h starting at HH.

The temperature background fields are interpolated vertically using routines from the HIRLAM preprocessing package for generating boundaries based on ECMWF model fields. These routines are based on the techniques for smooth transformation of large-scale analyses onto high-resolution grids described by Majewski (1985).

The vertical interpolation is carried out with an extra model layer inserted at the 2 m level. First a preliminary MESAN surface pressure field is estimated based on the MESAN orography together with the orography, surface pressure, temperature and humidity fields from HIRLAM. Pressures at the HIRLAM model levels above the high-resolution orography are then obtained from this preliminary surface pressure field. The planetary boundary layer (PBL) is treated separately during the vertical interpolation with the emphasis on preserving stability properties of the temperature profiles close to the surface. Above the PBL, interpolations are done using the logarithm of the pressure as the vertical coordinate. Inside the PBL, interpolations are done using a terrain-following vertical coordinate. To preserve stability in the PBL, the potential temperatures in the PBL are adjusted by a constant value so that they, at the top of the PBL (vertical ETA coordinate \( \eta = 0.8 \)), coincide with those from the interpolation in the free atmosphere. Finally the preliminary MESAN surface pressure is corrected in such a way that the geopotentials of pressure levels well above the highest mountains (500 hPa) equals the same geopotentials from the original HIRLAM fields. These corrections are generally quite small so no vertical re-interpolation is done to account for this change in the surface pressure.

The adjusted 2 m temperatures are then interpolated horizontally. The weight in the horizontal interpolation consists of two factors. One factor corresponds to the weight for a bilinear interpolation and the other factor varies depending on the difference in land fraction between the points on the background grid and the point for which the downscaling is calculated.

The first guess for the minimum/maximum 2 m temperature is calculated as the point-wise min/max of the 3-hourly 2 m temperature MESAN analyses. The analyses used for estimation of the minimum are from 1500 UTC (day before) to 0900 UTC. The estimate of the maximum is based on analyses from 0300 to 2100 UTC. The reason for these choices is to capture the extremes all over the domain covering different time zones corresponding to longitudes from 25°W to 45°E.

2.2.2. Wind

The surface wind depends heavily on the topography resolved at the scale of the NWP model. Moreover, observations are scarce and often not representative for the scale of NWP. Because of this, the 10 m winds were not analysed using MESAN. Instead the HIRLAM analyses at 0000, 0600, 1200, and 1800 UTC were downscaled to the higher-resolution MESAN grid, using the routines from the HIRLAM preprocessing package described earlier, and used as input to the dynamical adaptation, DYNAD.

In the dynamic adaptation the finer resolution results in new valleys opening and hill tops becoming higher which affects both the wind direction and velocity. The adaptation to these new relief characteristics is performed by running a simplified version of the HIRLAM model (DYNAD) on the fine-resolution grid. Since the time-scale of most physical processes relevant for NWP (e.g. condensation, precipitation formation, forcing by radiation, etc.) is much longer than that of dynamic adjustment of the wind, the only physical processes that are kept within DYNAD are vertical diffusion and wind advection. The lower boundary uses a weighted constant surface temperature and a weighted value of surface roughness. The wind is held constant at about 500 hPa during the adaptation period. Moreover, it is only necessary to integrate the model for a period of 30 min in order to arrive at an adaptation of the wind and mass fields. This means that the cost of running DYNAD is heavily reduced compared to running a full NWP model at the high resolution.

2.2.3. Precipitation

To accurately represent the hydrological cycle is a challenge for NWP models. Many parameters are involved and since...
cloud and precipitation measurements are not assimilated in HIRLAM, several parameters affecting precipitation are only constrained indirectly from other observations, e.g. temperature and humidity.

The daily precipitation (0600–0600 UTC) from HIRLAM is computed as the sum of two forecast differences: \((f_{o000+1800} - f_{c000+1800}) + (f_{c1200+1800} - f_{c1200+0600})\). The reason for doing this instead of simply using the \(f_{c000+2400}\) forecast is that the forecast model needs some time to spin up a realistic atmospheric state from its initial conditions. This state may not be in balance due to the lack of information about clouds and precipitation in the analysis. Several forecast differences were tested and the one above turned out to give the best result, probably because the analyses at 0000 and 1200 UTC are the ones based on the largest number of observations.

In EURO4M the ambition was to develop a downscaling for \(22\) based on wind speed and direction as well as topographical features of the finer-resolution grid (Johansson and Chen, 2003). Initially, attempts were made along these lines but with unsatisfactory results. In order to produce the first guess for the 2D analysis of daily precipitation within the project time limits, it was decided to settle for a bi-linear horizontal interpolation as the downscaling step. One may surmise that the EURO4M HIRLAM resolution of \(22\) km gives a good enough background for the interpolation to work reasonably well.

### 2.3. Surface analysis

For temperature and precipitation, the downscaling described in the previous section is then followed by an analysis step. MESAN is built on the Optimum Interpolation (OI) technique (Daley, 1991) where the analysis is a linear combination of the first guess and the innovations (differences between observations and the corresponding first-guess values). The weights in the linear combination depend on the assumed accuracy of the measurements as well as on the assumed quality of the first-guess field. Physiographic fields such as fraction of land/water, roughness length, topography are also included as input for the analysis.

#### 2.3.1. Data selection

In order to avoid the analysis drawing closer to any duplicated observations, a test is made on the station location. The priority order of the observations is of the ‘first entered, first served’ principle. Hence observations will be rejected if they are from locations too close to the ones that have already entered the system.

The fact that only a limited number of observations influences the analysis at a given grid point is used to speed up the calculations. However, in order to avoid discontinuities caused by different observations being used for nearby grid points, the box method for data selection is applied. This method uses two sets of boxes with different sizes. For all grid points in an analysis box (small), observations from a corresponding selection box (large) is used. The result is that, for the majority of two neighbouring analysis boxes, the used observation subset is identical.

#### 2.3.2. Structure functions

The spatial structure of the innovation vectors is incorporated into the first-guess error correlation functions in the OI method. MESAN uses empirical structure functions that are influenced by differences in land/sea fraction and differences in height between the analysed grid point and the grid point containing the observation:

\[
\rho(r, d_f, d_z) = \frac{1}{2} \left\{ e^{-\frac{r}{L}} + \left(1 + \frac{2r}{L}\right) e^{-\frac{r}{2L}} \right\} F_f(d_f)F_z(d_z).
\]

Here, \(r\) is the distance between two grid points, \(L\) is the characteristic horizontal scale (190 km for temperature and 270 km for precipitation). The functions \(F_f(d_f)\) and \(F_z(d_z)\) model the influence of differences in land-fraction, \(d_f\) and height, \(d_z\), between two points and vary linearly from 1 when \(d_f = d_z = 0\), to 0.5 for \(d_f = 1\) and \(d_z \geq 500\) m (Hågmark et al., 2000).

#### 2.3.3. Error specifications

In OI, as well as in variational data assimilation methods, the analysis-error variance is minimized. Therefore, the uncertainties in the first guess as well as in the observations need to be specified in terms of error standard deviations. The values used in the EURO4M 2D reanalysis are the same as those used in the operational runs with MESAN.

The first-guess error for the 2 m temperature is set to a constant. However different months are assigned different values. During the winter months (November–February) it is set to 8 K, during spring and autumn (August–October, March–April) it is 7 K, and during summer (June–July) it is 5 K. In retrospect, these numbers seem too high. Looking at the mean monthly standard deviation for the observation minus first guess in section 3.1, the value should probably be closer to 2 K during summer and 3 K during winter. However, aside from affecting the quality control, it is not the magnitude of, but the ratio between, the errors of the observations and the first guess that influence the analysis. For the minimum and maximum temperature, a similar scheme is used. Here the first-guess error is set to 6 K during winter (November–February), 5 K during March and October, 4 K for the summer (August–September) and 3 K for the remaining months. The observation error for \(2\) m in MESAN depends on the observed temperature (K) according to the following relation:

\[
\sigma_o = \begin{cases} 
1.5 & \text{if } T < 260, \\
0.15(260 - T) & \text{if } 260 \leq T < 270, \\
1.5 & \text{if } T \geq 270.
\end{cases}
\]

For the daily precipitation, the first-guess error is set to a constant of \(20\) mm day\(^{-1}\) while the observation error is given by

\[
\sigma_o = 0.7 + 0.2 \, rr,
\]

where \(rr\) is the observed daily rain rate (mm). Again, the value for the first-guess error seems too high. Considering the mean monthly standard deviation for the observation minus first guess in section 3.3, a value of about \(4\) mm day\(^{-1}\) seems more realistic.

#### 2.3.4. Quality control

Quality control is done according to a standard procedure (Lorenc, 1981). First a gross error check is done where observations are compared with the first guess. Then an analysis is performed and the observations are checked against this analysis. Observations are rejected if the difference between the observations and the first guess is greater than a limit based on the first guess and observation errors. The analysis is then redone for the boxes that contained rejected observations.

The limits used for the gross error check were the same as the ones used in the operational runs with MESAN at SMHI: \(5\sigma_o\) for the \(2\) m temperature, \(4\sigma_o\) for minimum and maximum temperatures and \(15\sigma_o\) for daily precipitation. Here \(\sigma_o\) refers to the standard error of the corresponding first guess.

#### 2.4. Observations

The strategy for adding detail with the surface 2D analysis is twofold. First, to downscale the HIRLAM 3D analysis using auxiliary high-resolution information about the orography and
the land/sea mask. Second, to use all available observations in the surface analysis. For example, the surface 2 m temperature analysis in HIRLAM only used analysis input data from ECMWF MARS archive (type=ai). Measurements of min/max 2 m and daily precipitation are not used at all in the HIRLAM analysis.

The source of observations for the 2D analysis depends on the parameter. Four databases have been used: the Integrated Surface Database (ISD; Smith et al., 2011), maintained by NOAA’s National Climatic Data Center (NCDC), the MARS archive of ECMWF, the European Climate Assessment & Dataset (ECA&D), including also the non-public observations used as input for E-OBS version 10.0 (Haylock et al., 2008), and the national climate databases of SMHI and Météo-France (MF).

As will be seen in the next section there are large areas of Europe that are poorly covered with observations. The major reason for this is the inability of many meteorological agencies and services to release their climate data to the public. Efforts have been made to open up the climate archives, both world-wide with WMO resolution 40 (WMO, 1995) and within Europe with Directive 2003/88/EC of the European Parliament and the Council on the Re-use of Public Sector Information (EP and Council, 2003). Still today many institutions in Europe impose conditions and charge for access to their climate data. This has the effect that, although the EURO4M analyses are freely available to the public, the observations used in the process are not.

Many observations appear in more than one of the data sources. The fact that MESAN keeps the first observation when presented with duplicates was used to prioritize between the different datasets. The contents of the ISD and ECMWF MARS archives are similar when it comes to 2 m temperature observations. However, ISD data come with a quality flag and it turned out that prioritizing the quality controlled data from ISD over observations from ECMWF MARS resulted in slightly lower analysis errors. The manually quality controlled content of the national climate archives of SMHI and MF is non-overlapping and was given highest priority.

Observations of 2 m were retrieved from SMHI, MF, ISH, and ECMWF MARS and were presented to the MESAN system in that order of priority. A map showing the station distribution for the different sources used in the 2 m analysis at 0000 UTC on 1 January 1989 is given in Figure 1(a). The vast majority of the approximately 2000 observations comes from ISD (or ECMWF). However, note the uneven distribution of available 2m data throughout Europe. Norway and Finland now have open data policies and so should be possible to cover better in future reanalyses.

Daily values with min and max of the 2 m temperature have been retrieved from SMHI, MF, and ECA&D in terms of the public as well as the non-public input data used for the production of E-OBS. Data from the national databases of SMHI and MF were considered to be of higher quality than the ECA&D data. The map showing the station distribution for the different sources used in the min and max analysis of 2 m at 1 January 1989 is shown in Figure 1(b). Again the distribution of the 4000 observations is uneven, with half of the available stations located in France.

Observations of daily precipitation come from the same sources as the min and max temperatures, i.e. from SMHI, MF, and ECA&D. Also here, data from the national databases of SMHI and MF were considered to be of higher quality than the ECA&D data. As an example, the map in Figure 1(c) shows the station distribution for the different sources used in the analysis of daily precipitation for 1 January 1989. The picture here is similar to that for the minimum temperature with some 50% of the stations being French.

3. Reanalysis performance

The increase in available reanalysis datasets has led to the growth of a diverse user base and their quality requirements have evolved accordingly. However, a basic requirement is that the reanalysis is in agreement with available observations.

The analysis system uses a downscaled HIRLAM forecast as a first guess. The departures between the first guess and the observations contain information about the quality of the reanalysis, and its evolution in time. Analysis departures contain information about how closely the reanalysis fits the observations. However, the degree of this fit is to a large extent controlled by the prescribed first-guess and observation-error statistics. Normally the analysis departures are not so useful for measuring quality since they only monitor the fit to dependent data. Here, the analysis departures are instead calculated using independent data from a ten-fold cross-validation.

Another source of information about the performance of the analysis system is the analysis increments, i.e. the difference between the first guess and the resulting analysis. However, one needs to be careful with the interpretation. Small increments can be the sign of the model producing a good first guess, but they can also result from a lack of observations. Systematic increments often stem from residual biases in observations and/or the forecast model. The analysis increments can also be used to spot errors in observations, e.g. concerning station positions or reports where missing data have been replaced with zeros. One such example can be seen in the analysis increments in section 3.1 where the ‘hot spot’ northwest of the British Isles was found to be associated with the weather station Alna at 10.83°W, 59.93°N. Since only observations over land have been used in the 2D analysis, this
seemed suspicious. Alna is actually located at 10.83° E in the Oslo municipality, Norway, and the error was tracked down to incorrect coordinates for Alna in the ISD database during 2007–2010.

In the following sections the MESAN first guess and analysis, together with forecasts and analyses from HIRLAM, are compared to observations. Here, a ten-fold cross-validation technique was used. The MESAN analysis was redone ten times; each time every tenth observation was excluded and used for the error calculations. After ten passes, errors were calculated at all observation points using independent data. Only observations that were not used by HIRLAM are included in comparisons involving HIRLAM analyses.

The bias is calculated as the average difference between modelled (forecast or analysis) values and observations. Hence, a negative/positive bias means that the model gives too low/high values.

The significance of any added value is also checked. To test whether the error of one model is stochastically lower than the other, the Mann–Whitney–Wilcoxon (MWW) test is employed (Gibbons and Chakraborti, 2010). The reason for this choice is that neither the RMSE nor the absolute error follows a normal distribution. Still, the MWW test is almost as efficient as a t-test when applied to normal distributions.

The idea behind the MWW test is that, given two samples X and Y (of size n1 and n2), the ranks of the Xs in an ordered combination of the two samples would generally be larger than the ranks of the Y’s if the values of the X population are statistically larger than those of the Y population. The test statistic \( U \) is based on the sum of ranks of the Xs and Ys within this combined ordered set. In our application X and Y correspond to analysis error samples of equal size and hence \( n1 = n2 \).

### 3.1. Temperature

The fit to observations for the 2m background, first guess and analyses during the years 1989–2010 is illustrated in Figure 2(a). It shows monthly means of the bias and standard deviation at 0000, 0600, 1200 and 1800 UTC for HIRLAM forecasts (H22), down-scaled HIRLAM forecasts (FG05) and MESAN analyses (M05) together with the number of observations used in the validation (N). Mean monthly values for all years are shown in Figure 2(b). Here comparisons are included also for HIRLAM forecasts (BG05) and HIRLAM analyses (AN05), both corrected for the difference in height between the 22 and 5 km grid. The correction is made by assuming a constant lapse rate of 6.4 K km\(^{-1}\).

The downscaling (FG05) produces better results than the horizontally interpolated background from HIRLAM (H22) throughout the time period, but is only slightly better than a simple downscaling based on the constant-lapse-rate assumption (BG05). Hence, the main reason why the MESAN analyses (M05) perform better than the HIRLAM analyses, compensated for height differences (AN05), is probably due to the use of more information about some station locations was changed for the ECMWF MARS archive, but is not yet fully understood. Perhaps the information about some station locations was changed for this time period. If the change in position was large enough, these observations would no longer be regarded as duplicates and so would enter the analysis.

The performance of the analyses of minimum and maximum 2 m temperatures is illustrated in Figure 4. Monthly means of the bias and standard deviation of the daily differences between observations and the first guess (fg) and the analysis (an) are...
presented together with the number of observations used for the analysis in Figure 4(a) for the minimum temperature and in Figure 4(b) for the maximum temperature. The analyses draw towards observations and reduce the bias to almost zero. As expected, the first guess (point-wise minimum/maximum from the 3-hourly MESAN t2m analyses corresponding to night/day) has problems capturing the extreme temperatures and is hence too warm when predicting the minimum and too cold when predicting the maximum temperatures. The annual cycles can again be explained by the synoptic variability being larger during wintertime. The standard deviation for the first guess and analysis departures are of similar order for both the minimum and maximum temperatures, while the bias is about twice as high for the maximum temperature. There seems to be a slight trend caused by the increase in the quality of the MESAN analyses used for the first guess, in turn (both min and max). The reason for this may be an increase in the number of observations used for the first guess, in turn caused by the increase in t2m observations over time (Figure 2(a)).

For the minimum temperature (\(n_m\)), the median absolute errors for the first guess and the analysis are 0.81 and 0.61 K. The corresponding values for the maximum temperature (\(tx\)) are 0.91 and 0.51 K. The error distributions differ significantly in the two cases (MWW, \(tn: U = 982, n1 = n2 = 31 366 874, P < 0.01\) two tailed. \(tx: U = 2132, n1 = n2 = 36 859 197, P < 0.01\) two tailed).

3.2. Wind

The wind climate in Europe depends to a large extent on topography and the paths of low pressure systems. The 2D downscaling and analysis of the 10 m wind does not contribute any information to the 3D analysis over the sea but provide some detail in mountainous areas.

Wind observations are not used in the 2D analysis (nor in HIRLAM or DYNAD) so here the comparison can be done against independent data without resorting to cross-validation. The evolution of the monthly means of the error standard deviation and bias for HIRLAM and DYNAD are shown in Figure 5(a). Here the departures are calculated relative to observations of the 10 m wind speed retrieved from the MARS archive at ECMWF. Mean values for each month of the year, averaged over the whole period is given in Figure 5(b).

The larger synoptic variability during winter is also reflected in the performance of the wind analysis where both models have larger errors during wintertime.

Looking at Figure 5(a), the performances of HIRLAM and DYNAD are almost identical. The step-like change in bias for both models during the years 1993–1997 is probably related to some errors/differences in the observations (the wind bias for ERA-Interim show the same pattern for this time period). DYNAD appears to add some value to the HIRLAM analysis during the winter and spring (Figure 5(b)).

When checking for added value, the distance between the modelled and observed wind vectors was used as an error measure in order to incorporate both magnitude and direction into one number. The median distance error is then 2.07 m s\(^{-1}\) for HIRLAM and 2.06 m s\(^{-1}\) for DYNAD (MWW \(U = 66.2, n1 = n2 = 31 257 882, P < 0.01\) two tailed). Note that the difference, although significant in a statistical sense, is negligible.

One reason for the small impact could be that the errors are calculated for the whole domain while the dynamic downscaling only makes a difference in areas with steep terrain. To check this,
a separate validation was done over Norway (strong orography) for the year 2010. For this case the median distance error was 2.46 m s\(^{-1}\) for HIRLAM and 2.44 m s\(^{-1}\) for DYNAD (MWW \(U = 3.14, n_1 = n_2 = 126 180, P < 0.01\) two tailed), again a negligible difference. Separating the error into speed and direction it turns out that the wind speed is of equal quality while the wind direction is better in DYNAD (at about 3/4 of the 214 stations). Further investigation is needed to find out if this is due to tunnelling effects caused by the high-resolution orography in DYNAD.

The geographical variation in the dynamical downscaling with DYNAD is shown in Figure 6 where the norm of the average wind vector increment (HIRLAM minus DYNAD) is used as the measure. First one can note that the downscaling of the wind makes a difference only in high-terrain areas. The wind analysis increments depend on a combination of orographic and synoptic effects. During the winter there is more synoptic activity in northwestern Europe and its combination with high orography (e.g. in Norway) is illustrated in Figure 6(a). The synoptic activity is shifted towards the southern and southeastern parts of Europe during summer as shown in Figure 6(b) where the Atlas mountains and the area south of the Black Sea stand out.

3.3. Daily precipitation

As noted earlier, it is not an easy task to get a good representation of the hydrological cycle. Since HIRLAM is not assimilating rain rate measurements, the precipitation produced by the model is based on information about surface pressure, temperature and humidity derived from the assimilated observations. Imbalances in the analysed fields relative to the model equations can result in spin-up/spin-down effects. To avoid this, only forecasts longer than 6 h were used to produce the first guess.

Since the downscaling of the daily precipitation from HIRLAM consists only of horizontal interpolation, there is no difference between the information in the background from HIRLAM and that in the first guess to MESAN. All the added value in the MESAN analysis comes from the information in the observations.

The fit to observations is illustrated in Figure 7. Figure 7(a) shows monthly means of the bias and standard deviation for the HIRLAM forecasts and the MESAN analyses together with the number of observations used in the validation (N). Mean monthly values for all years are presented in Figure 7(b). The median absolute error for rainy events (\(rr24 > 0.1\) mm) is 1.6 mm for HIRLAM and 0.67 mm for MESAN. The distributions for the absolute error differ significantly (MWW \(U = 2791, n_1 = n_2 = 39 788 077, P < 0.01\) two tailed).

HIRLAM has very little overall bias when it comes to daily precipitation during both January and July as shown in Figure 7(b). However, looking at maps with analysis increments for these two months, it turns out that HIRLAM is too wet in both the Scandinavian mountains and the Alps in January as well as in July (Figure 8). During July, HIRLAM has a wet bias throughout most of central and eastern Europe.

© 2016 The Authors. Quarterly Journal of the Royal Meteorological Society published by John Wiley & Sons Ltd on behalf of the Royal Meteorological Society.
4. Extreme weather events

Extreme weather may result in loss of human lives and property. The economic losses due to such events have, according to IPCC (Handmer et al., 2012), been estimated to range from a few US$ billion to above 200 billion (since 1980, evaluated at 2010 prices).

To see how the 2D analysis of 2 m temperature, 10 m wind, and daily precipitation represent extreme weather events, four cases were selected: the European heat wave during the summer 2003, the European cold wave during the winter 2009–2010, the cyclones Lothar and Martin in 1999, and the 2002 European floods.

Note that the datasets shown in the figures of this section may contain values outside the range indicated by the colour bars. The range has been chosen in order highlight differences in the data. Values below or above the limits are shown with the colours of the limit values.

4.1. Heat wave 2003

The 2003 European heat wave was the hottest summer on record in continental Europe since at least 1540 and temperatures rose above 40 °C in several locations. In Belgium, France, Germany, Italy, the Netherlands, Portugal, Spain, Switzerland and the United Kingdom, more than 66 000 deaths were attributed to the heatwaves. In the European Alps, the average thickness loss of glaciers reached the equivalent of about 3 m of water, nearly twice as much as during the previous record year of 1998 (WMO, 2011).

In Figure 9 the focus is on an area around France and northern Italy which was at the centre of the event. The detail added by the 2D MESAN analysis is clearly seen by comparing Figure 9(a) and (b). Even if one corrects the HIRLAM 2 m temperature for the height differences between its orography and the one used in MESAN, with a fixed lapse rate adjustment, there is still added value in the MESAN analysis. This is shown in Figure 9(c) where the difference between the RMSE for HIRLAM (H22LR, corrected for height differences) and MESAN (M05) turn out positive at most locations. Here, the median RMSE for HIRLAM and MESAN is 1.2 and 0.87 K respectively; the distribution of the errors differ significantly (MWW \( U = 12.6, n_1 = n_2 = 1906, P < 0.01 \) two tailed).

In the area shown, the 99%iles (≈ 6500 cases) for the HIRLAM analyses (corrected for height differences), the MESAN analyses and the observations during the heat wave are 36.3, 36.1, and 36.2 °C respectively. These numbers suggest that both HIRLAM and MESAN represent the warm observations rather well.

4.2. Cold wave 2010

The European cold wave during the 2009–2010 winter was part of a larger pattern with extremely low temperatures over large parts of the Northern Hemisphere. In continental Europe, the episode resulted in over 450 deaths (WMO, 2013).
By contrast there were very mild conditions over the Arctic and Canada. These conditions were associated with large-scale atmospheric disturbances connected to the Arctic and North Atlantic Oscillations and the El Niño event (WMO, 2011).

For this case the focus is on Scandinavia where differences between HIRLAM and MESAN can be seen in high-terrain areas. However, the cold wave also affected areas with flat terrain in this region, e.g. most parts of southern Sweden, where the temperature never rose above 0 °C during the entire month of January.

The magnitude of the event is similar in both HIRLAM and MESAN as can be seen in Figure 10(a, b). However, the details are better described by MESAN, even when the HIRLAM temperatures are corrected with a fixed-lapse-rate adjustment. Looking at the difference between the RMSE for HIRLAM (H22LR, corrected according to fixed lapse rate) and MESAN (M05), HIRLAM has a larger error at a vast majority of the observation stations (Figure 10(c)). The median RMSE for HIRLAM and MESAN is 1.1 and 0.83 K respectively; the distribution of the errors differ significantly (MWW \( U = 11.7, n_1 = n_2 = 1338, P < 0.01 \) two tailed).

The extreme temperatures at the tail of the distribution (99%iles, \( \approx 5800 \) cases) are again rather similar in HIRLAM (compensated for height differences) and MESAN: \(-26.3 \) and \(-25.8 \) °C respectively, both being higher than the corresponding value for the observations: \(-27.2 \) °C.
4.3. Cyclones 1999

During 36 h over 26–28 December 1999, two successive extreme wind storms, named Lothar and Martin, hit western Europe leading to major damage and many human fatalities. Lothar crossed and affected several countries (France, Germany, Switzerland and Italy) while the effects of Martin were more localized to southwestern France. Both storms were associated with a very intense upper-level zonal jet and very strong baroclinicity (Riviere et al., 2010).

The maps in Figure 11(a, b) show the point-wise maximum 10 m wind speed from the analyses at 0000, 0600, 1200, and 1800 UTC during 26–28 December 1999 from HIRLAM and DYNAD respectively. DYNAD provides more detail and enhances the wind in high-terrain areas. A comparison with observed maximum wind speeds over France from the same dates and synoptic times shows that the mean RMSE is only slightly smaller for HIRLAM: 3.3 m s\(^{-1}\) compared to 3.4 m s\(^{-1}\) for DYNAD (no significant difference). The spatial distribution of the errors is given in Figure 11(c). However, the median absolute errors for HIRLAM and DYNAD are 2.0 and 1.8 m s\(^{-1}\) when all observations during the 36 h are used in the comparison. In this case, the error distributions differ significantly (MWW\(U = 3.83, n_1 = n_2 = 3064, P<0.01\) two tailed).

To get an idea of the dynamical downscaling on the extreme values of the maximum 10 m wind during these storms, one can compare the 95th percentiles (approx 30 cases). The result is 18.7 m s\(^{-1}\) for HIRLAM, 19.2 m s\(^{-1}\) for DYNAD and 23.0 m s\(^{-1}\) for the observations. Again, the performance is very similar.

4.4. Floods 2002

The meteorological situation leading to the August 2002 floodings was marked by two low pressure systems travelling from the western Mediterranean basin north-eastwards. This atmospheric circulation pattern is known to be related to flooding in central and eastern Europe (Mudelsee et al., 2004). The result in 2002 was flooding of historic proportions in central Europe (including Germany, the Czech Republic, Austria, Romania, Slovakia). More than 100 people were killed and more than 450 000 people had to be evacuated. The damage was estimated to US$ 9 billion in Germany alone (WMO, 2011).

For this event HIRLAM produces too much precipitation, as can be seen by comparing its accumulated amounts for August 2002 in Figure 12(a) with the corresponding amounts from MESAN in Figure 12(b). This result is consistent with the wet HIRLAM summer bias shown in Figure 8(b). To illustrate the added value in the MESAN analysis, the difference in RMSE of daily precipitation during August 2002 is presented in Figure 12(c). Here, the difference is once again in favour of MESAN, with a median RMSE equal to 0.72 mm compared to 1.3 mm for HIRLAM. The error distributions differ at the 1% level of significance (MWW\(U = 35.3, n_1 = n_2 = 2079\)).

During the month of August 2002, the 99th percentiles of daily precipitation from the HIRLAM forecasts, the MESAN analyses and the observations were 29.1, 28.7 and 28.9 mm day\(^{-1}\) respectively. These numbers suggest that HIRLAM as well as MESAN represents the extreme precipitation rather well during this event.

5. Conclusions and outlook

This article documents the configuration and performance of the EURO4M 2D reanalysis of the 2 m temperature, its daily minimum and maximum, the 10 m wind, and the daily precipitation over the past two decades. For these parameters, this is the most complete homogeneous dataset available for Europe today.

It has been demonstrated that the EURO4M 2D reanalysis has added value over the EURO4M 3D reanalysis which in turn has added value over the ERA-Interim reanalysis. The latter was shown in Part 1 of this article. The verification was done against independent observations by means of a ten-fold cross-validation and the Mann–Whitney–Wilcoxon test was used to test for statistical significance. The added value in terms of lower median absolute or RMS errors was shown to be significant for all parameters in the 2D reanalysis. However, although significant, the added value by the DYNAD adaptation of the 10 m wind is very small. The success of DYNAD is limited to situations
with a quasi-steadiness of the PBL and does not work well in cases of inversions, frontal passages, and locally induced thermal circulation (Zagar and Rakovec, 1999).

Analysis increments were studied to see where and how MESAN and DYNAV adds information to the first guess from HIRLAM. This showed that during winter, HIRLAM is generally too cold at night and too warm during the day. The wind analysis is actually a dynamical adaptation of the HIRLAM wind (no observations used) and the result is a combination of orographic and synoptic effects. During the winter there is more synoptic activity in northwestern Europe while the synoptic activity is shifted towards the southern and southeastern parts of Europe during summer. For the precipitation it turns out that HIRLAM is too wet in both the Scandinavian mountains and the Alps in January as well as in July. There is also a wet bias throughout most of central and eastern Europe during July.

In order to see how the reanalysis performs in extreme situations, a number of cases were selected. To summarize, the patterns look realistic and MESAN/DYNAV adds value to the HIRLAM analysis (lower RMSE) and both HIRLAM and MESAN/DYNAV produce reasonable extreme values for the 2003 heat wave and the 2002 floods. The extreme values for the 2009–2010 cold wave and the 1999 cyclones are somewhat underestimated.

All in all, the 2D reanalysis has been shown to produce realistic results and could be useful in the evaluation and further development of climate models, as a part the European reference material for analysis of climate change and as a basis for decision-makers and for climate adaptation measures. The fundamental limitation on our ability to analyse the state of the atmosphere lies in the availability of observations with good quality. Despite efforts by the EU and WMO, lots of observations are still withheld by available parameters and download alternatives. The EURO4M 2D reanalysis dataset is freely available and most of it is published by John Wiley & Sons Ltd on behalf of the Royal Meteorological Society.

Acknowledgements

The project ’European Reanalysis and Observations for Monitoring’ (EURO4M) was funded by the European Commission under the EU Seventh Research Framework Programme theme 9 ’Space’, grant agreement no. 242093. The authors would like to thank the reviewers whose constructive comments and suggestions substantially improved the main text. Special thanks go to Elise van den Besselaar and the people at the ECA&D group, KNMI, for providing the E-OBS input files and to François Besson at Météo-France for preparing the French observations. Thanks also to Sébastien Villeaume for help with the archiving of this dataset in the MARS system at the National Supercomputer Centre (NSC) in Linköping. Per Källberg and Per Undén, our colleges at SMHI, have also contributed to the work with valuable comments and fruitful discussions.

References

Daley R. 1991. Atmospheric Data Analysis. Cambridge University Press: Cambridge, UK.

Dee D, Balmaseda M, Balsamo G, Engelen R, Simmons A, Thepaut J-N. 2014. Toward a consistent reanalysis of the climate system. Bull. Am. Meteorol. Soc. 95: 1235–1248, doi: 10.1175/BAMS-D-13-00043.1.

Desroziers G, Berre L, Chapnik B, Poli P. 2005. Diagnosis of observation, background and analysis error statistics in observation space. Q. J. R. Meteorol. Soc. 131: 3363–3365, doi: 10.1256/qj.05.035.

EP Council. 2003. Directive 2003/98/EC of 17 November 2003 on the re-use of public sector information. Off. J. Eur. Union 46: 90–97.

Gibbons JD, Chakaraborti S. 2010. Nonparametric Statistical Inference (5th edn). CRC Press: Boca Raton, FL.

Haggmark I, Jarsson KI, Golivik S, Olofsson PO. 2000. MESAN, an operational mesoscale analysis system. Tellus 52A: 2–20, doi: 10.3402/tellusa.v52i1.15236.

Handmer J, Honda Y, Kundezwicz ZW, Arnell N, Benito G, Hatfield J, Mohamed IF, Peduzzi P, Wu S, Shrestyukov B, Takahashi K, Yan Z. 2012. Changes in impacts of climate extremes: human systems and ecosystems. In Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation, Field CB, Barros V, Stocker TF, Qin D, Dokken DJ, Ebi KL, Mastrandrea MD, Mach KJ, Plattner GRK, Allen SK, Tignor M, Midgley PM (eds.): 231–290. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press: Cambridge, UK, and New York, NY.

Haylock MR, Hofszt NA, Klein Tank AMG, Klok EJ, Jones PD, New M. 2008. A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. J. Geophys. Res. 113: D20119, doi: 10.1029/2008JD010201.

Johansson B, Chen D. 2003. The influence of wind and topography on precipitation distribution in Sweden: Statistical analysis and modelling. Int. J. Climatol. 23: 1523–1535, doi: 10.1002/joc.951.

Lorenz AC. 1981. A global three-dimensional variational statistical analysis scheme. Mon. Weather Rev. 109: 701–721.

Majewski D. 1985. Balanced initial and boundary values for a limited-area model. Beitr. Phys. Atmos. 58: 147–159.

Mudelsee M, Börngen M, Tetzlaff G, Grunewald U. 2004. Extreme floods in central Europe over the past 500 years: Role of cyclone pathway ’Zugstrasse Vb’. J. Geophys. Res. 109: D23101, doi: 10.1029/2004JD005034.

Riviere G, Arbogast P, Maynard K, Ioly A. 2010. The essential ingredients leading to the explosive growth stage of the European wind storm Lothar of Christmas 1999. Q. J. R. Meteorol. Soc. 136: 638–652, doi: 10.1002/qj.565.

Smith A, Lott N, Vose R. 2011. The integrated surface database: Recent developments and partnerships. Bull. Am. Meteorol. Soc. 92: 704–708, doi: 10.1175/2011BAMS3015.1.

WMO. 1995. ’WMO resolution 40 (Cg-XII): WMO policy and practice for the exchange of meteorological and related data and products including guidelines on relationships in commercial meteorological activities’. http://www.wmo.int/pages/about/Resolution40_en.html (accessed 6 Octo- ber 2014).

WMO. 2011. ’Weather extremes in a changing climate: Hindsight on foresight WMO No. 1075’. http://www.wmo.int/pages/media/ancetre/news/documents/1075_en.pdf (accessed 26 March 2015).

WMO. 2013. ’The global climate 2001–2010 a decade of climate extremes – summary report, WMO No. 1119’. http://library.wmo.int/pmb/chief/wmo_1119_en.pdf (accessed 26 March 2015).

Zagar M, Rakovec J. 1999. Small-scale surface wind prediction using dynamic adaptation. Tellus 51A: 489–504, doi: 10.1034/j.1600-0870.1999.00141.x.

Dataset access

The EURO4M 2D reanalysis dataset is freely available and most of it is published as a CLIPC pilot project activity via the Earth System CoG at https://www.earthsystemcog.org/projects/cog/ (accessed 25 April 2016). Click on a Federated ESFG-CoG node and enter ‘mesan’ in the field ’Search & Download Data’ to get a list of available parameters and download alternatives.

© 2016 The Authors. Quarterly Journal of the Royal Meteorological Society published by John Wiley & Sons Ltd on behalf of the Royal Meteorological Society.

Q. J. R. Meteorol. Soc. 142: 2132–2146 (2016)