

RESEARCH ARTICLE

CERRA, the Copernicus European Regional Reanalysis system

Martin Ridal¹  | Eric Bazile² | Patrick Le Moigne² |
Roger Randriamampianina³ | Semjon Schimanke¹ | Ulf Andrae¹ |
Lars Berggren¹ | Pierre Brousseau² | Per Dahlgren³ | Lisette Edvinsson¹ |
Adam El-Said² | Michael Ginton² | Susanna Hagelin¹ | Susanna Hopsch¹ |
Ludvig Isaksson¹ | Paulo Medeiros¹ | Esbjörn Olsson¹ | Per Unden¹ | Zheng Qi Wang³

¹Swedish Meteorological and Hydrological Institute, Norrköping, Sweden

²Meteo France, Toulouse, France

³Norwegian Meteorological Institute, Oslo, Norway

Correspondence

Martin Ridal, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden.

Email: martin.ridal@smhi.se

Funding information

Copernicus Europe,
Grant/Award Number: C3S_322Lot1

Abstract

A regional reanalysis has been produced for a domain covering entire Europe from 1984 to 2021. The reanalysis is produced as part of the Copernicus Climate Change Service. The Service provides the high-resolution deterministic Copernicus European Regional Reanalysis (CERRA), run at a horizontal resolution of 5.5 km, a 10-member ensemble run at 11-km resolution as well as an offline surface analysis, CERRA-Land. CERRA-EDA uses an ensemble data assimilation (EDA) technique to perturb the initial condition of the different members. Apart from the horizontal resolution the CERRA and CERRA-EDA setups are the same; for example, the same data assimilation scheme, same physics parameterization as well as the same vertical levels. These new systems are built from HARMONIE cy40 version, including some back-phased physics from a newer model version (cy42). Conventional observations, satellite-based radiances, atmospheric motion vector winds and bending angle from radio occultation observations are used. In addition, ground-based zenith total delay (ZTD) from global navigational satellite systems (GNSS) and local surface observations, rescued from historical archives at the local National Meteorological Services, are used. Another new feature is the construction of the background error statistics for the data assimilation. Information from the ensemble run, CERRA-EDA, is used in the derivation of the background error statistics for the high-resolution CERRA runs. These background error statistics are updated every second day. By doing so, daily environment variation is taken into account as well as all variations over the 37 years of production. The reanalyses and reforecasts from CERRA show an added value compared to the global ERA5 for almost all variables at the surface level. This becomes particularly clear when selecting smaller areas with complex terrain where the high resolution is beneficial. In the free atmosphere it is primarily the analyses and short forecasts, 3–6 hours, that give an added value.

KEYWORDS

Copernicus climate change service, data assimilation, regional reanalysis

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Authors. *Quarterly Journal of the Royal Meteorological Society* published by John Wiley & Sons Ltd on behalf of the Royal Meteorological Society.

1 | INTRODUCTION

Reanalyses are more and more widely used in atmospheric sciences. As compared to long observation series and climate simulations, a reanalysis optimally combines observations and models, using state-of-the-art numerical weather prediction (NWP) models. This provides a spatially complete and physically consistent dataset of the best estimate of the state of the Earth's atmosphere (Dee et al., 2014). Reanalysis does not have the constraint of issuing timely forecasts, so there is more time to collect observations, and when going further back in time, to allow for the ingestion of improved, reprocessed versions of the original observations. By doing so, all available observations will be used and, if possible, with better quality than the original ones. This will all benefit the quality of the reanalysis product.

In this paper a regional reanalysis is presented that covers the whole of Europe for the time period 1984–2021. This new reanalysis is run with a high horizontal and temporal resolution compared to the existing global reanalyses. Global reanalyses have been run for quite some time, for example the ERA-40 (Uppala et al., 2005) and ERA-Interim (Dee et al., 2011) produced by the European Centre for Medium-Range Weather Forecasts (ECMWF), the Japanese JRA-55 (Kobayashi et al., 2015), the American Modern-Era Retrospective Analysis for Research and Applications (MERRA) (Rienecker et al., 2011), and the NCEP/NCAR global reanalysis (Kalnay et al., 1996). These were typically run with a horizontal resolution of 60–70 km or lower. The most recent reanalysis from ECMWF, ERA5, is run with a horizontal resolution of 31 km (Hersbach et al., 2020).

As smaller-scale information and more realistic representation of extreme events are more and more demanded by users, the interest for spatially highly resolved regional reanalyses (RRAs) is increasing.

Examples of such RRAs are the early North American Regional Reanalysis (Mesinger et al., 2006) and the Arctic System Reanalysis (Bromwich et al., 2012, 2017) followed by recent efforts in Europe like the Met Éirann reanalysis (MÉRA) at 2.5 km covering Ireland and the British Isles (Gleeson et al., 2017) and the Copernicus Arctic Regional ReAnalysis (CARRA) covering parts of the Arctic region (Køltzow et al., 2022).

Within the European Union-funded project European Reanalysis and Observations for Monitoring (EURO4M), the work started to answer the need for regional reanalyses (RRAs) covering Europe with a high resolution compared to what the global RA can provide. EURO4M delivered RRA products at an intermediately high resolution (22 km) and also with downscaling and surface analysis versions at higher resolutions (Dahlgren et al., 2016).

EURO4M was followed by the FP7 (7th Framework Project for Research) project “Uncertainties in Ensembles of Regional Re-Analyses” (UERRA, www.uerra.eu). UERRA increased the resolution to address some limitations of EURO4M and also focused on the uncertainties in the reanalyses. The time period of the RRA in UERRA was also much longer than in EURO4M. In order to assess the uncertainties in the RRA, advanced ensemble data assimilation (EDA) was used for a long time period. High-resolution deterministic RRA and other gridded datasets were also included in the evaluation of the uncertainties.

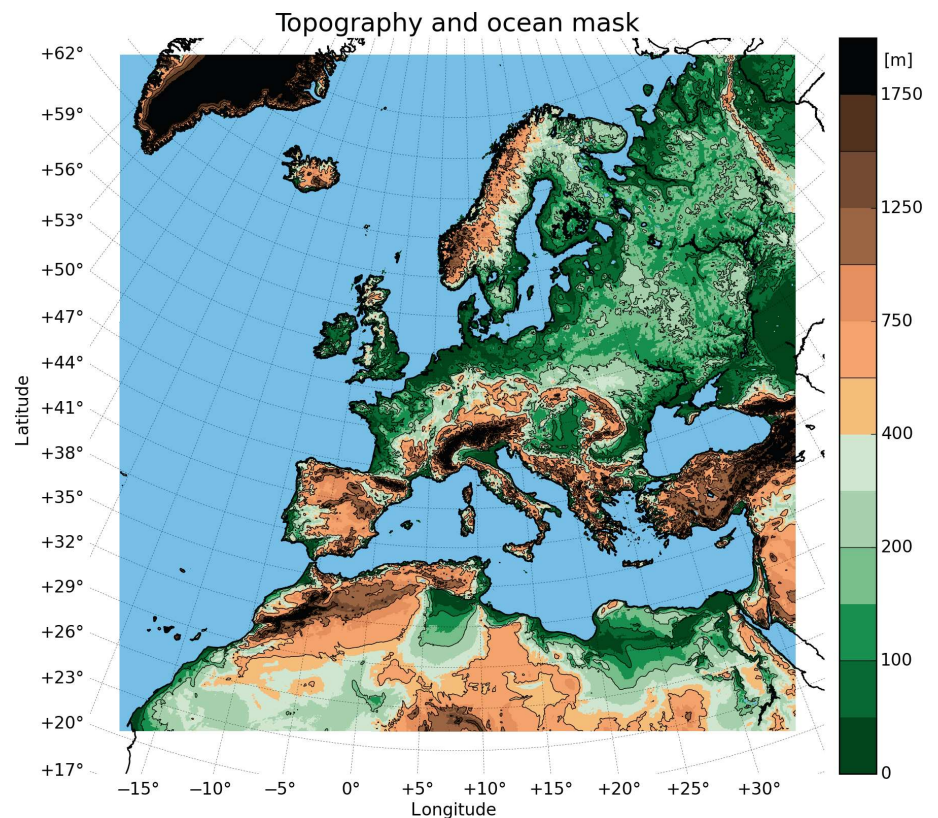
Within the framework of UERRA a regional reanalysis was produced using the HARMONIE (HIRLAM ALADIN Regional/Mesoscale Operational NWP In Europe) system. HARMONIE is a complete system for NWP. It was developed in the HIRLAM (High Resolution Limited Area Model) consortium and builds upon the code of the models ALADIN (Aire Limitée Adaptation Dynamique Développement International), AROME (Applications of Research to Operations at Mesoscale) and ALARO (ALADIN and AROME combined model) developed in a collaboration between Météo-France and the ALADIN and HIRLAM consortia.

As part of the UERRA project two other reanalyses were also produced, the COSMO-REA6 run by DWD (Bollmeyer et al., 2015) and the UM-RRA from UK Met Office. These RRAs were run with different systems and will not be described further here.

The long reanalysis using the HARMONIE system within the UERRA project (UERRA-HARMONIE) was run from 1961 to 2015 with a horizontal resolution of 11 km and the ALADIN physics parameterization. Atmospheric and surface data assimilation were both included. To introduce large-scale information from the global reanalyses a large-scale constraint has been added to the cost function. Two reports presenting the results from the UERRA reanalysis are available at <http://www.uerra.eu/> as UERRA deliverables D2.6 (Ridal et al., 2016) and D2.7 (Ridal et al., 2017), respectively. While the first report describes preliminary results, the second one contains results from the entire period from 1961 to 2015.

As part of the Copernicus Climate Change Services, the UERRA-HARMONIE system was used to extend the time series for the period 2016–July 2019. The entire dataset, starting in 1961, is available in the Copernicus Climate Data Store (CDS). Meanwhile, the system for the next generation regional reanalysis for Europe was developed. The new system, CERRA (Copernicus European Regional ReAnalysis), was developed with a horizontal resolution of 5.5 km that covers entire Europe (Figure 1). Together with the high resolution this new system contains many improvements compared to UERRA-HARMONIE. It is

FIGURE 1 The CERRA domain including topography. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/qj.4764)]



for example based on a newer version of HARMONIE and includes many more observations, where the biggest difference is the inclusion of satellite observations. Moreover, CERRA is run with a three-hour cycling whereas UERRA-HARMONIE was run with a six-hour cycling frequency. The number of vertical levels is also increased, together with a higher model top in order to better describe the processes in the upper troposphere and lower stratosphere.

In parallel to the high-resolution deterministic system, an ensemble system of 10 members with a lower resolution has been developed. This system uses an ensemble data assimilation (EDA) system, referred to as CERRA-EDA, that runs with 11-km horizontal resolution. Although CERRA-EDA uses a six-hour cycling frequency, its setup is very similar to that of CERRA with, among others, the same number of vertical levels and the same observations (apart from one satellite instrument).

One innovative feature of the CERRA reanalysis is that the output from CERRA-EDA is used to create its background error statistics (El-Said et al., 2022). This means that CERRA-EDA needs to run a few days ahead of the deterministic system. More details are provided in the following sections.

In addition to CERRA and CERRA-EDA a two-dimensional surface analysis was produced, CERRA-Land (Verrelle et al., 2022). This analysis utilizes additional observation such as carefully quality-controlled

precipitation observations. CERRA-Land will be described in a separate paper and not further discussed here. However, for most of the surface and soil variables it is recommended for users to obtain these from CERRA-Land instead of CERRA.

The need for keeping the European regional reanalysis up to date and improved with a more modern system and higher resolution is motivated by the high number of users. At the end of the contract UERRA-HARMONIE had more than 1500 registered users and CERRA and CERRA-Land data was downloaded by more than 600 users during a period of one year (August 2022 to August 2023). The users come from a wide range of research areas, for example, the breeding range and time of migrating song birds (Vega et al., 2021) to other applications, for example the planning of new wind farms (Spanghel et al., 2023).

In parallel to CERRA a regional reanalysis for the Arctic has also been produced. The Copernicus Arctic Regional ReAnalysis (CARRA) is run over two smaller domains with a resolution of 2.5 km (Køltzow et al., 2022). Since the two systems, CERRA and CARRA, have been developed and produced in parallel it has been possible to take advantage of many synergies between the two projects. For example, many local observations could be used by both systems as well as some monitoring tools.

In this paper details of the modeling systems are provided and results from CERRA are compared to the global ERA5 reanalysis which is used as lateral

boundaries. Furthermore, ERA5 is considered to be the state-of-the-art reanalysis at the moment and is in some senses more advanced than CERRA; for example, ERA5 is using a more advanced data assimilation scheme, although run at a much lower horizontal resolution of 31 km. Running with higher resolution is one of the main advantages of CERRA and the purpose of producing a regional reanalysis. CERRA also uses a few additional observation sources compared to ERA5.

The paper is organized as follows; Section 2 describes the model system setup, including the ensemble setup and the construction of the background error statistics. Section 3 describes the observations used while Section 4 shows a few results as compared to ERA5, including a case study of a storm that hit Scandinavia in 2005. Section 5 shortly gives some detail on the technical implementation followed by some conclusions in Section 6.

2 | THE CERRA MODEL SYSTEM

HARMONIE is a complete system for NWP. It was developed in the HIRLAM consortium and builds around the code of the models ALADIN, AROME and ALARO developed in collaboration with Météo-France and the consortia ALADIN and HIRLAM (Bengtsson et al., 2017; Termonia et al., 2018).

The CERRA reanalysis system builds in large parts on HARMONIE cy40h1.2 with some important differences. Among others, several changes in the script system compared to the reference version of HARMONIE were introduced to speed up the production. The main achievement was to separate the analysis and forecast steps. In the production runs, the new analysis is started as soon as the first guess is available, that is, the three-hour forecast. The remaining forecast hours are run in parallel to the next analysis. This saves a lot of time in a reanalysis setup but is of no use for operational forecasts. Several other optimizations have been introduced in the CERRA system to improve the performance. These include, among others, a few code and script changes, including changes to the data handling; for example, fetching observations for one month at a time or several days of boundary files, and changes to the submission commands by altering the number of nodes or memory allocations for different tasks.

The CERRA system runs with a three-hour cycling, 5.5-km horizontal grid distance over the same domain as used for UERRA-HARMONIE (Figure 1). It is run with 106 vertical levels with the lowest level at 10 m, 19 levels below 1000 m and the model top at 1 hPa. In addition, a 10-member ensemble, CERRA-EDA, described

in Section 2.4, is run in parallel to CERRA at 11-km horizontal resolution. CERRA uses a time step of 4 min while CERRA-EDA runs with a 10-min time step.

The main focus of CERRA is to optimize the short-range forecasts, that is, the 3–12-hour forecasts. The three-hour forecast is the background for the next analysis and for longer forecast lead times the influence from the boundaries will increase in a limited-area model. Forecasts up to 30 hours lead times are, however, provided to users for continuity purposes, for example, for daily mean calculations.

2.1 | Improved physics

Another difference compared to HARMONIE cy40h1.2 is that parts of the ARPEGE-ALADIN physics package have been back-phased from cy42t2 used in the Météo-France ARPEGE operational system in order to improve primarily convective precipitation at 5.5 km. This ARPEGE physics is used operationally at Météo-France in the global system since summer 2019 with a horizontal resolution of 5.5 km over France.

The radiative transfer in the atmosphere with the surface is described using the RRTM scheme (Rapid Radiative Transfer Model) for long-wave radiation (Mlawer et al., 1997) and the six-band Fouquart–Morcrette scheme for short-wave radiation (Fouquart & Bonnel, 1980; Morcrette, 1991). Gravity waves, their reflection and trapping, as well as upstream blocking, are taken into account (Catry et al., 2008).

The boundary layer parameterization is represented by a mass-flux shallow convection scheme using a CAPE closure (Bechtold et al., 2001) and by a diffusion scheme based on prognostic turbulent kinetic energy (Cuxart et al., 2000) using the Bougeault and Lacarrere (1989) mixing length modified by the shallow convection (Bazile et al., 2011). Deep convection is represented with a mass-flux scheme based on a moisture convergence closure (Bougeault, 1985) modified by Marquet et al. (2019) to avoid grid point storms. The microphysical processes are based on Lopez (2002) and a statistical cloud scheme of Smith (1990) modified by Bouteloup et al. (2005).

2.2 | Data assimilation

Observations are introduced into the model through data assimilation, both in the three-dimensional (3D) atmosphere and in the two dimensional surface scheme. The assimilation scheme used for the atmospheric analyses is a 3D-Var assimilation scheme, which creates an analysis by minimizing a cost function (e.g., Gustafsson et al., 2001;

Lindskog et al., 2001; Fischer et al., 2005; Brousseau et al., 2008):

$$J(x) = \frac{1}{2}(x - x_b)^T \mathbf{B}^{-1}(x - x_b) + \frac{1}{2}(y - H(x))^T \mathbf{R}^{-1}(y - H(x)), \quad (1)$$

where x is the model state vector (containing the control variables vorticity, divergence, temperature, specific humidity and surface pressure), x_b is the first guess or background, in our case a three-hour forecast. y represents the observations while H is the observation operator, $\mathbf{B}$ is a matrix that describes the errors of x_b and $\mathbf{R}$ is a matrix that describes the errors of the observations y . It is assumed that the observation errors are spatially uncorrelated and thus, $\mathbf{R}$ is represented as a diagonal matrix. The background error matrix, on the other hand, describes both spatial correlations and balances between variables.

CERRA-EDA uses a background error correlation ($\mathbf{B}$) matrix usually constructed from a mix of forecast differences from summer and winter to get an estimate valid for the whole year (Berre, 2000; Montmerle & Berre, 2010; Brousseau et al., 2012). For CERRA analyses, however, $\mathbf{B}$ is constructed using statistics from the CERRA-EDA output. The methodology of $\mathbf{B}$ matrix construction is described in Section 2.4.

The surface observations are assimilated using an optimal interpolation (OI) method using CANARI (Code for the Analysis Necessary for ARPEGE for its Rejects and its Initialization) and SURFEX (surface externalisée) analysis schemes.

CANARI (Taillefer, 2002) is a part of the IFS/ARPEGE (Integrated Forecast System/Action de Recherche Petite Echelle Grande Echelle) source code and was developed to provide both surface and atmospheric ARPEGE/ALADIN analysis based on the OI method. In CERRA, while using the surface part of CANARI, for the 2-m temperature and 2-m humidity analysis, the MESCAN option, which takes into account the difference in height and the land-sea mask, was used.

The surface analysis is performed in two steps. First CANARI finds the analysis increments in each grid point based on observations minus first guess. In the next step a consistent update of the SURFEX fields (surface and deep soil temperature, soil moisture and snow depth) is made based on analysis increments interpolated to all grid points by CANARI.

2.3 | Geographical data

Geographical data are all the external data used as input of the system to describe as realistically as possible the physiographical data at any location in the domain. It includes topography, land use, soil texture, and so forth.

2.3.1 | Land use dataset ECOCLIMAP

Land use is described in CERRA by the ECOCLIMAP database (Faroux et al., 2013). The ECOCLIMAP program is a dual database at 1-km resolution that includes an ecosystem classification and a coherent set of land surface parameters that are primarily mandatory in meteorological modeling (notably leaf area index and albedo). Hence, the aim of this physiography is to enhance the quality of initialization of the surface parameters and impose some surface attributes within the scope of weather forecasting studies, including reanalysis. Fractions of nature, sea, urban areas and inland water are calculated by ECOCLIMAP. The fraction of land equals one for pixels covered with vegetation only and decreases for those where urban fraction, inland water fraction and/or sea fraction are positive. In CERRA the urban model is not activated since at the start of the project it was not fully developed and validated. Instead the urban fractions were converted to rocks which are part of the natural areas. This is a common approach in, for example, climate models.

In large lakes such as Ladoga or Onega, the lake fraction equals 1 in the middle of the lakes indicating pure lake pixels whereas on their border the lake fraction decreases to the benefit of the vegetation fraction. This is also the case for smaller lakes where the lake fraction never reaches 1. Changes in land use are not considered so that the prescribed land use is constant over the entire reanalysis period.

2.3.2 | Topography and soil texture

Two other important pieces of information are required by the system. These are the topography and the soil texture for natural tiles. The topography is used to describe the elevation of each grid point, to account for pressure effects on flows and which is for instance used in the conversion from air temperature to potential air temperature. Topography is computed from the Global Multi-resolution Terrain Elevation Data (GMTED2010) (Danielson & Gesch, 2011).

In the ISBA (Interaction Soil–Biosphere–Atmosphere) soil–vegetation–atmosphere transfer model some parameters depend on soil texture which is represented by the fractions of sand and clay, especially pedotransfer functions. These fractions are derived from the Harmonized World Soil Database (HWSD, Nachtergaele et al., 2012) that is a soil map at 1-km resolution that combines several datasets available worldwide.

2.3.3 | Lakes

In CERRA the lake parameterization model FLAKE is used (Kourzeneva, 2010). The lake depth comes from the

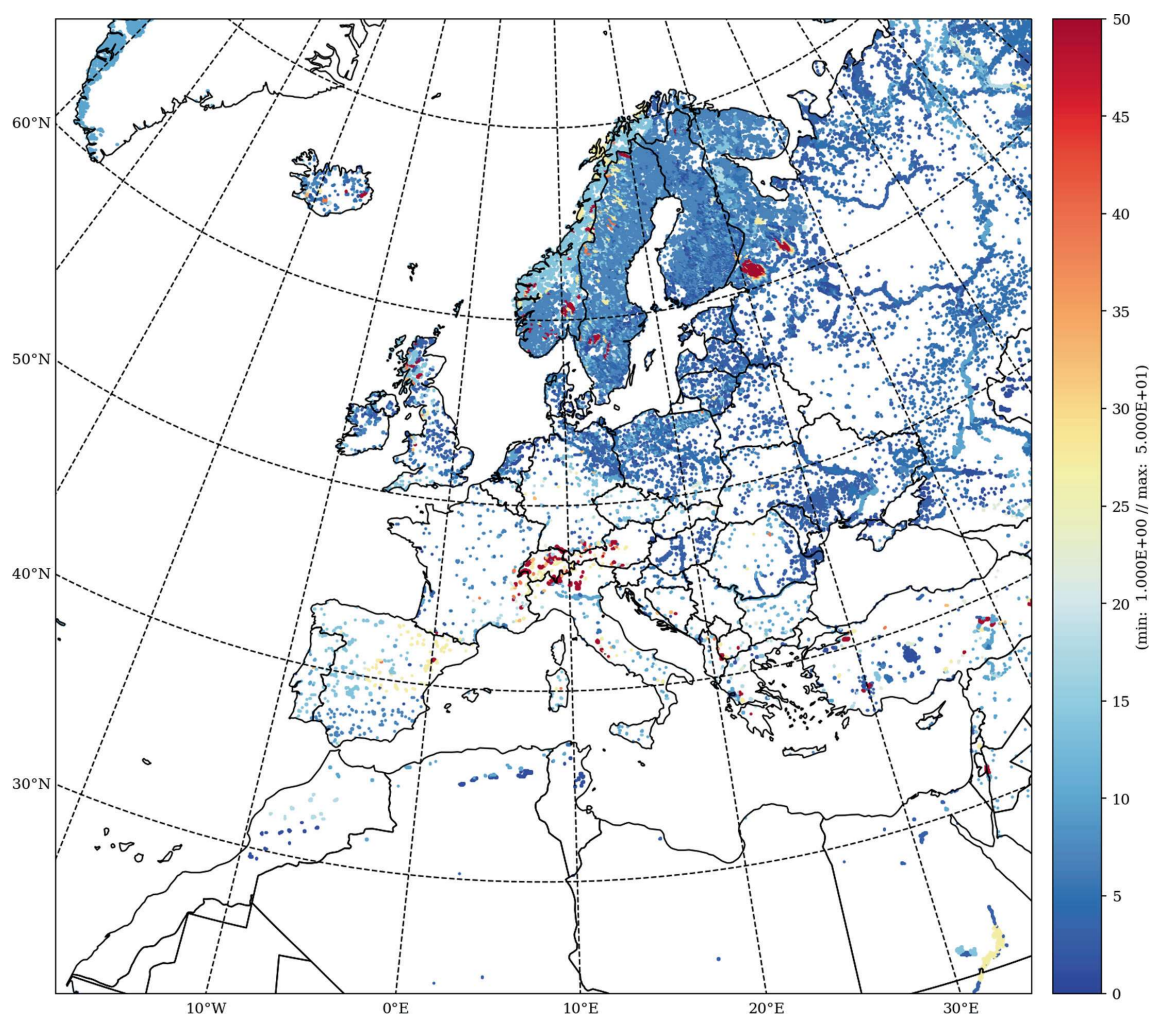


FIGURE 2 Lake depth in CERRA. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/qj.4764)]

Global Lake DataBase (GLDB, version 2) which is a 1-km global lake depths database. Figure 2 shows the distribution of lakes and their depths in CERRA. The high density of lakes in the northern part of the domain demonstrates the importance of considering lakes in the system.

2.3.4 | SST and sea ice

Sea surface temperatures (SST) and sea ice concentrations are taken from different sources depending on the period and the location. ERA5 data (prior to 1985) and OSTIA (Operational Sea Surface Temperature and Ice Analysis, <https://ghrsst-pp.metoffice.gov.uk/ostia-website/index.html>) (Donlon et al., 2012) products are used for the Mediterranean Sea, the Black Sea, the Atlantic and the North Sea. For the Baltic Sea, an ocean reanalysis based on the RCO model (Rossby Centre Ocean model) is used, which is available for the period 1970–1999 (Liu et al., 2013). Between the year 2000–2016 a Baltic

Sea reanalysis based on the NEMO-Nordic model (Hordeir et al., 2019) was applied with data available through Copernicus Marine Services (CMEMS, <https://marine.copernicus.eu/>). After 2016, data based on the Danish North Sea—Baltic Sea model HBM (<http://ocean.dmi.dk/models/hbm.uk.php>) was used for the Baltic Sea (Poulsen & Berg, 2012). This dataset was also available through CMEMS. The availability of this dataset ceased in September 2020 so from October 2020 and onward OSTIA data were used also for the Baltic Sea.

2.4 | The ensemble setup

The CERRA-EDA system is an Ensemble of Data Assimilation system consisting of 10 members. It is built from the same code as the deterministic (CERRA) version and uses the same observations, except the MSU (Microwave Sounding Unit) satellite observations from the early years. There are a few differences where the main ones are that

the EDA system runs with an 11-km horizontal resolution and cycles every six hours. This also means that the time step can be longer, 10 min instead of 4 min as for CERRA.

The distinction between each ensemble member is obtained by perturbing the observations prior to the data assimilation. One control member is initialized with unperturbed observations while the other nine are initialized with perturbed observations. The perturbation technique uses a random technique assuming a Gaussian distribution within the estimated observations errors. There are no perturbations to the SST/sea ice or the physical parameters.

As lateral boundary conditions for CERRA-EDA the ERA5 ensemble is used. The ERA5 EDA system has 10 members and is run with 63-km horizontal grid distance (Hersbach et al., 2020).

The resulting analyses are used to produce six-hour forecasts from which the background error covariances are estimated. Due to the fact that CERRA uses the output from CERRA-EDA, the EDA system needs to run ahead of the deterministic system. This is normally not a problem since CERRA-EDA is faster than CERRA due to the lower resolution and the longer time step. The 10 members run more or less in parallel so this does not slow the production down much either.

The CERRA-EDA derived **B** matrices comprise two principal components, a seasonal component and a daily component. The seasonal component is derived from two one-month experiments using 5.5 km resolution and a 4-min time step. One experiment is for a winter month, January 2018, and one from a summer month, July 2017. This referred to the “static” or “climatological” **B** in this paper and is the **B** matrix used in the CERRA-EDA runs. Within CERRA however, this is done a bit more advanced (El-Said et al., 2022).

First a seasonal component from these two months of forecasts is introduced. This seasonal component represents seasonal changes by incorporating forecast differences coming from each season, summer or winter, according to the current time of year. During winter the majority of the forecast differences will come from the winter period and then gradually change so that in summer the majority will come from the summer month.

Secondly a day-to-day variability is introduced using forecast differences from CERRA-EDA, interpolated to 5.5 km. The daily component comprises a 2.5-day moving average, in parallel to CERRA production. The choice of averaging over 2.5 days was made by studying the day-to-day variability of the prognostic variables. It was found that the maximum gradient in the background error estimate, particularly for vorticity, occurred over 2–3 days. The **B** matrix statistics for CERRA are updated every second day.

The pooling of forecast differences is such that 80% come from the seasonal component and 20% from the daily. The reason for choosing an 80%–20% seasonal–daily split for EDA forecast differences was that they provided the most favorable root-mean-squared errors (RMSE) forecast differences when compared to a static seasonal-only **B** matrix. Forecast scores and data assimilation performance statistics as well as a more detailed description of the **B** matrix construction can be found in El-Said et al. (2022).

3 | OBSERVATION USAGE

Different sources of observations have been used in the CERRA and CERRA-EDA production runs, such as conventional and satellite observations as well as other non-conventional observations. These are described further below.

3.1 | Conventional observations

Conventional observations include in-situ data from land surface stations (SYNOP), surface measurements from ships (SHIP), drifting buoys (DRIBU), radiosonde/balloon profiles (TEMP, PILOT) and aircraft measurements (AIREP, AMDAR, ACARS).

The atmospheric data assimilation in CERRA and CERRA-EDA utilizes the conventional observations that are currently stored in the Meteorological Archival and Retrieval System (MARS) archive at ECMWF. Since ECMWF has produced several reanalysis products their archive of conventional observations is fairly well maintained for the CERRA domain during the production period from 1984 until present.

Although comprehensive, the archive at ECMWF does not include every observation that has been made in the time period from 1984 until present. For CERRA, local SYNOP data from the archives at the Swedish, Norwegian, Danish, Finnish and Icelandic meteorological institutes are added on top of the data retrieved from MARS at ECMWF. Additional surface observations from the climate database at Météo-France are also used. Some of the latter observations have already been used in the UERRA-HARMONIE and MESCAN-SURFEX reanalyses, but additional observations from automatic stations are available for use in CERRA.

Figure 3 shows the total number of conventional observations per day used in the CERRA reanalysis. The data are smoothed with a 30-day running average to be more clear.

There are a few interesting things to notice in Figure 3. First, the number of radiosondes (TEMP) keeps increasing until around 2010. From there on to the end of the

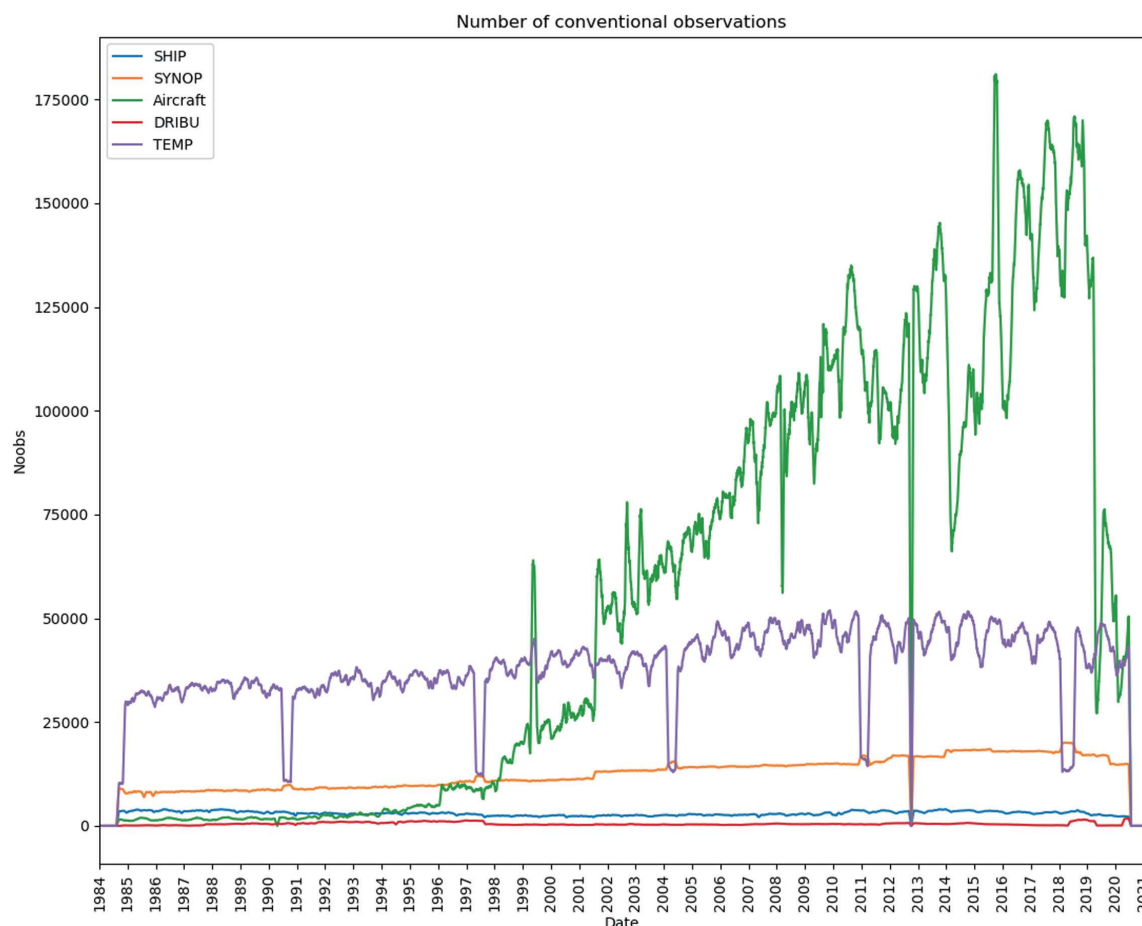


FIGURE 3 Number of daily observations used in the CERRA production showing ship observations (SHIP), land surface stations (SYNOP), aircraft observations (AIRCRAFT), drifting buoys (DRIBU), and radiosonde observations (TEMP). The data are smoothed over 30 days for display purposes. [Colour figure can be viewed at wileyonlinelibrary.com]

period, a constant number, or even a small decrease can be observed. The reason is that radiosondes are expensive to launch and the evolution of other observations, for example, aircraft observations, reduces the need for radiosondes. Several institutes have therefore reduced the number of radiosonde launches. At the beginning of the production, that is, the start of the five first streams, some radiosonde data were missing. This was quickly fixed and for the rest of the reanalysis period all data are present. Another interesting thing is the steep drop of aircraft observations at the end of the reanalysis period. This is caused by the reduced number of aircrafts due to the Covid-19 pandemic.

3.2 | Satellite radiances

The satellite radiances used in the reanalysis are primarily from the microwave sounders MSU, AMSU-A and AMSU-B (Advanced MSU). The latter is also referred to as MHS (Microwave Humidity Sounder). These instruments

have been flown since 1980 on a number of platforms measuring temperature and humidity that are very useful for NWP. In Table 1 the channels used from each instrument are summarized as well as the platform it was mounted on and the time period from which we use the data.

For the later period of the reanalysis, after 2007, radiances from the Infrared Atmospheric Sounding Interferometer (IASI) are also utilized. From this hyperspectral instrument, a subset of 74 channels is used to derive temperature and moisture at different levels of the atmosphere.

Included in Table 1 are also the variational bias correction predictors used for each particular instrument. In some cases it can differ slightly between different channels for the same instrument depending on what altitude the channel is most sensitive to but in general they are the same. The bias correction method and the meaning of the different predictors is described in Section 3.4. An evaluation of the satellite radiance impact on the analysis and forecasts in CERRA can be found in Wang and Randriamampianina (2021).

TABLE 1 Summary of the usage of satellite radiances including the variational bias correction predictions used.

Satellite	Channels	Predictors used	Time period
Instrument: MSU			
NOAA-6	2–4	0, 1, 2, 8, 9, 10	1980–1983
NOAA-7	2–4	0, 1, 2, 8, 9, 10	1981–1985
NOAA-9	2–4	0, 1, 2, 8, 9, 10	1985–1987
NOAA-10	2–4	0, 1, 2, 8, 9, 10	1988–1992
NOAA-11	2–4	0, 1, 2, 8, 9, 10	1989–1995
NOAA-12	2–4	0, 1, 2, 8, 9, 10	1993–1997
NOAA-14	2–4	0, 1, 2, 8, 9, 10	1996–2007
Instrument: AMSU-A			
NOAA-15	5–11, (12)	0, 1, 2, 8–10, 15–18	1999–2021
NOAA-16	5–11, (12)	0, 1, 2, 8, 9, 10	2001–2009
NOAA-17	5–11, (12)	0, 1, 2, 8, 9, 10	2002–2003
NOAA-18	5–11, (12)	0, 1, 2, 8–10, 15–18	2006–2021
NOAA-19	5–11, (12)	0, 1, 8, 9, 10	2008–2021
Metop-A	5–11, (12)	0, 1, 8, 9, 10	2007–2021
Metop-B	5–11, (12)	0, 1, 8, 9, 10	2013–2021
Instrument: AMSU-B			
NOAA-15	3–5		
NOAA-16	3–5	0, 1, 8, 9, 10	2001–2007
NOAA-17	3–5	0, 1, 8, 9, 10	2002–2009
Instrument: MHS			
NOAA-18	3–5	0, 1, 8, 9, 10	2006–2021
NOAA-19	3–5	0, 1, 8, 9, 10	2008–2021
Metop-A	3–5	0, 1, 8, 9, 10	2007–2021
Metop-B	3–5	0, 1, 8, 9, 10	2013–2021
Instrument: IASI			
Metop-A	74 channels	0, 1, 2, 8, 9, 10	2007–2021
Metop-B	74 channels	0, 1, 2, 8, 9, 10	2013–2021

Note: The predictors are described in Section 3.4.2 and listed in Table 3.

Figure 4 shows the same as Figure 3 but for satellite radiances. Interesting to note in Figure 4 is the total dominance of IASI observations after they became available in 2007. IASI is a multispectral instrument and even if we only use 74 channels in the data assimilation it is dominating in the number of observations. Interesting to see is also that MSU data are available from the start of the reanalysis but gradually decrease as the instrument is phased out. After 2007, MSU is fully replaced by the AMSU instruments. Other things to notice are that AMSU-B is gradually replaced by MHS, which is basically the same instrument and that AMSU-A is gradually decreasing in number toward the end of the period. Data from the

METOP-C satellite became available in spring 2021. They are not used in this reanalysis but should be considered in a continuation or a next generation of reanalysis.

3.3 | Other non-conventional observations

Apart from satellite radiances a number of non-conventional observation types are included. Several of these observations are obtained from reprocessed datasets. Only the last few years are real-time or near-real-time observations. This means that the quality

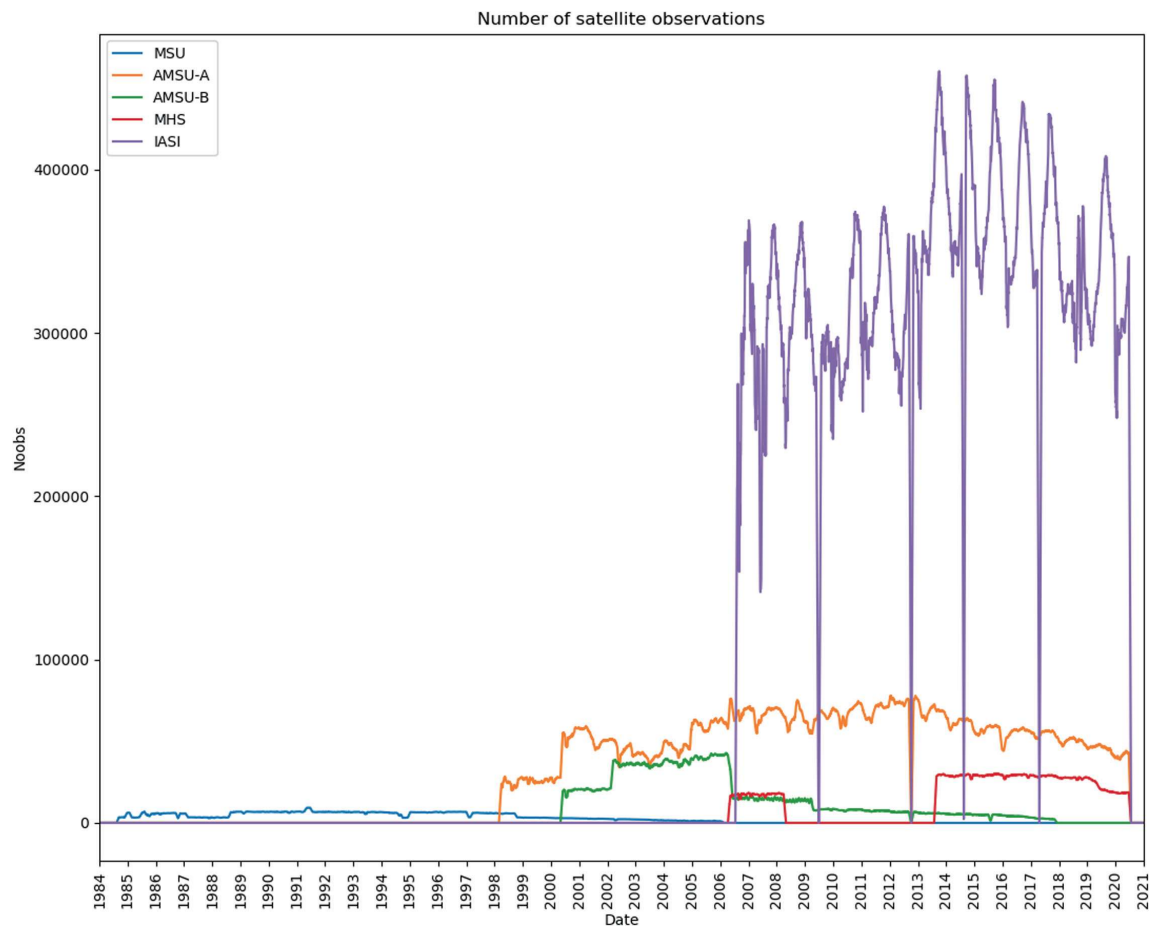


FIGURE 4 Same as in Figure 3 but for satellite radiances from MSU, AMSU-A, AMSU-B, MHS and IASI. [Colour figure can be viewed at wileyonlinelibrary.com]

can be regarded as very good and that the signal-to-noise ratio in many cases has been improved. It also means that the blacklisting, for example, minimum altitude to use or similar, can be a bit more generous than in operational model configurations.

A reprocessed dataset of zenith total delay (ZTD) from ground-based global navigation satellite systems (GBGNSS) from the EPN EUREF Permanent Network (Pacione et al., 2017) is used for the CERRA reanalyses. The ZTD basically depends on the density of the atmosphere and the moisture content. If the ground pressure is known a good estimate of the integrated water vapor of the atmosphere can be made (Sanchez Arriola et al., 2016; Lindskog et al., 2017). ZTD data are available from 1997 onward and there are up to 342 stations in this dataset but there are not data from all of them at any one time and stations may come and go during the period. All 342 stations are allowed in the system through a white-list, thereby allowing any of them to be used when they appear.

The refraction, or bending angle, of the GNSS signal by the atmosphere observed by a receiver satellite is known as GNSS radio occultation (GNSS-RO). These observations

give very accurate profiles of temperature, humidity and pressure (e.g., Poli et al., 2010 or Bauer et al., 2014). The radio occultation meteorology satellite application facility (ROM-SAF) provides a reprocessed GNSS-RO dataset, available from 2001. The advantage of using a reprocessed dataset is that the signal-to-noise ratio is increased and the observations can be used closer to the ground than in the operational dataset. To be on the safe side the observations are used down to 3 km in CERRA even though the data are said to be of good quality all the way to the ground.

Another advantage of the GNSS-RO observations is that they are considered to be bias-free. This means that they can be used as an anchoring observation during the bias correction of other observations, that is, help to improve the impact of other observations.

Atmospheric Motion Vectors (AMV) from geostationary and polar orbiting satellites have successfully been assimilated by the HARMONIE 3D-Var analysis system. For reanalysis purposes historical AMV data from different sources have been archived at ECMWF. Data from both geostationary and polar orbiters are available over the CERRA domain back to the early 1980's.

Observations of surface winds over sea can be obtained from scattermeters. In CERRA and CERRA-EDA a reprocessed dataset obtained from the ocean and sea ice SAF (OSI-SAF) at the Royal Netherlands Meteorological Institute (KNMI) has been used. In this dataset observations are available from 1997 to 2014 from the satellites ERS1 and -2, QuickSCAT, METOP-A and Oceansat-2. After 2014 near-real-time (NRT) data from ASCAT on METOP-A and -B were used. A detailed description of the processing and assimilation of scatterometer data in HARMONIE can be found in Valkonen et al., 2017.

Figure 5 shows the usage of the other non-conventional observations in the same way as in Figures 3 and 4. Noticeable here is that the number of observations is very dependent on the satellite constellation at the time. It can therefore seem a bit jumpy. Note also that AMVs are available from the start of the period with a drastic increase with better coverage from the polar orbiting satellites from about the year 2000.

The ground-based GNSS observations are treated as a subtype of SYNOP in the model. They are therefore not

included in Figure 5 but instead as part of the number of SYNOP observations in Figure 3.

Table 2 summarizes the datasets that are used in the reanalysis from different non-conventional observations other than satellite radiances.

3.4 | Variational bias correction

All observations contain some sort of observation error. These can, in most cases, be divided into two parts, one random error and one systematic error. The random errors, or uncertainties in the observations, are described in the observation error matrix used in the minimization and compared to the background errors. The systematic error, or bias, on the other hand, may not be a true error but rather the systematic differences between model and observations. These are harder to describe but may be compensated for using a bias correction procedure. This correction can either be static, that is, all observations from the same instrument are “corrected” with a constant

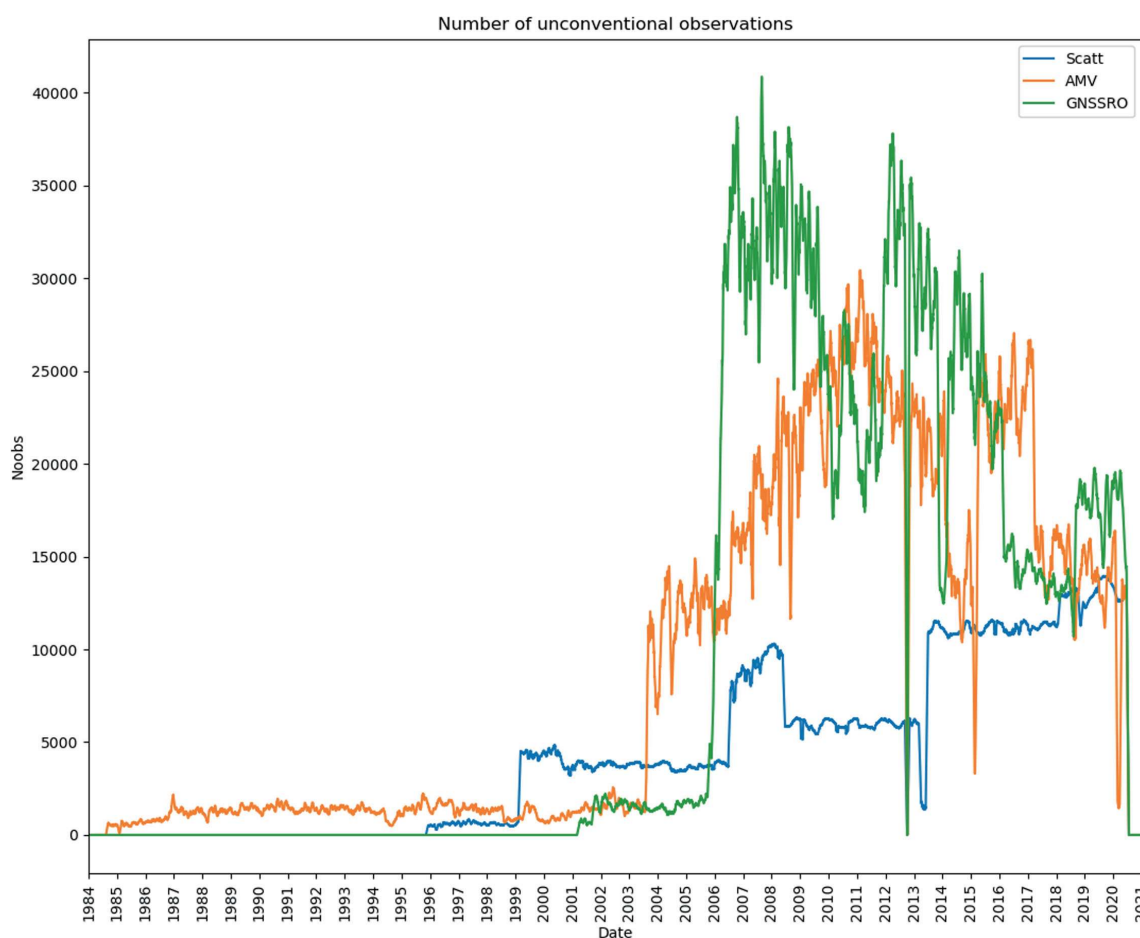


FIGURE 5 Same as in Figure 3 but for scatterometer winds, atmospheric motion vectors (AMV) and global navigational satellite systems (GNSS) radio occultation (RO) observations. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Summary of non-conventional observations other than satellite radiances.

Ground-based GNSS		
Stations	Platform	Time period
342 stations	Surface	2000–2021
Radio occultation GNSS		
Instrument	Platform	Time period
CHAMP	CHAMP	2001–2009
GRACE	GRACE-A, -B	2007–2011
COSMIC	FORMOSAT-3	2007–2011
GRAS	METOP-A, -B	2007–2021
AMV		
Instrument	Platform	Time period
AVHRR	NOAA-satellites	1980–2021
MODIS	TERRA	1999–2021
MODIS	AQUA	2002–2021
AVHRR	METOP-A	2007–2021
AVHRR	METOP-B	2012–2021
SEVIRI	MFG (Geostationary)	1982–2007
SEVIRI	MSG (Geostationary)	2002–2021
Scatterometer winds		
Instrument	Platform	Time period
ERS-1, -2	Scatterometer	1997–2001
SeaWinds	QuikSCAT/seawinds	1999–2009
ASCAT	METOP-A	2007–2021
OSCAT	Oceansat-2	2009–2021

value, or by a variational bias correction (VARBC). The latter is important if the bias is slowly varying, for example between day and night, different seasons or depending on the weather itself. How the bias correction varies is determined by a number of predictors valid for the specific observation type.

Traditionally mainly remote-sensing observations have been bias-corrected since these often are indirect observations that need to be processed to get an estimate of, for example, temperature from satellite radiances (Auligné et al., 2007) or moisture from ground-based global navigational satellite system (GNSS) observations (Sanchez Arriola et al., 2016). Another contributing factor is the model top. Some of the applied predictors (see Section 3.4.2) cannot be used with a relatively low model top, or badly represented stratospheric layers with too low resolution, for example.

3.4.1 | Method

The method used is the same as introduced by Derber and Wu (1998) at National Centers for Environmental Prediction (NCEP) and later by Dee (2004) at ECMWF. The bias b in each observation, radiance channel or observation station, is determined according to:

$$b(x, \beta) = \sum_{i=0}^{N_p} \beta_i p_i(x), \quad (2)$$

where p_i is the predictor i and β_i are unknown bias parameters. N_p is the total number of predictors used for the specific observation. This bias b is then subtracted from the observations in the second term in the cost function, described in Section 2.2, and the bias correction coefficient, β , is estimated by introducing an extra term in the cost function:

$$J(x) = \frac{1}{2}(x - x_b)^T \mathbf{B}^{-1}(x - x_b) + \frac{1}{2}(\beta - \beta_b)^T \mathbf{B}_\beta^{-1}(\beta - \beta_b) + \frac{1}{2}(y - H(x) - b(x, \beta))^T R^{-1}(y - H(x) - b(x, \beta)), \quad (3)$$

where $\mathbf{B}_\beta$ is the error correlation matrix for the bias correction coefficient. This basically only determines how quickly the bias correction coefficients are allowed to respond to the latest observations. The values have been determined by tuning in previous studies.

3.4.2 | Predictors

The bias between model and the observations can be caused by many different things. Several parameters can therefore be used as predictors. Examples of such predictors can be surface pressure, the thickness of different layers in the atmosphere or the state of the instrument itself (e.g., the different viewing angles for the different fields of view). Table 3 lists the predictors available in the CERRA system.

Different satellite channels from the same instrument can also have different predictors depending on which altitude the channel is more sensitive to, for example, predictor 1 (in Table 3) for low-peaking channels or predictor 2 (in Table 3) for high-peaking channels.

3.4.3 | Observations to be corrected

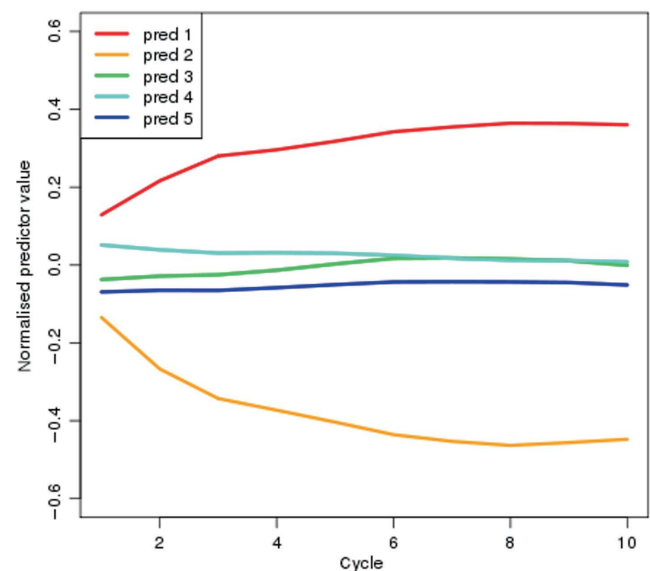
In the HARMONIE system a static bias correction can be applied to radiosonde temperatures which rely on tables containing vendor -specific information, station identifier,

TABLE 3 Predictors available for variational bias correction in the CERRA system.

Predictor number	Predictor
0	1 (constant)
1	1000–300 hPa thickness
2	200–50 hPa thickness
3	Skin temperature
4	Total column precipitable water
5	Stratosphere 1, 2–10 hPa thickness
6	Stratosphere 2
7	Surface wind speed
8	Satellite nadir viewing angle
9	(Satellite nadir viewing angle) ²
10	(Satellite nadir viewing angle) ³
11	(Satellite nadir viewing angle) ⁴
12	Cosine solar zenith angle
13	Not assigned
14	TMI diurnal bias
15	0 over sea, 1 over land
16	0 over sea, nadir viewing angle over land
17	0 over sea, (nadir viewing angle) ² over land
18	0 over sea, (nadir viewing angle) ³ over land
19	$\ln(\text{rain rate}+1)$
20	$\ln(\text{rain rate}+1)^2$
21	$\ln(\text{rain rate}+1)^3$

solar elevation and pressure level. This set of static corrections for reanalysis was derived at the University of Vienna during the ERA-CLIM2 project, led by ECMWF. The same corrections are used in ERA5. These corrections ensure that the bias-corrected time series of radiosonde temperatures are temporally homogeneous (Haimberger et al., 2017). The corrections are stored in MARS in ODB2 (observation data base) format and the procedure to use them involves a retrieval and then a pre-processing of the ODB2 file into a binary file which is then read by the model and applied to the observations.

For the satellite radiances used in CERRA, described in Section 3.2, the variational bias correction is very similar. For most channels predictors 0, 1, 8, 9 and 10, according to Table 3, are used with the exception of a few high-peaking channels where predictor 1 is omitted. Since the model top is fairly high for CERRA, 1 hPa, predictor 2 was included for the same channels. This predictor is not used in the operational HARMONIE configuration due to a lower model top, so extensive testing was performed prior to the production started (see also Wang & Randriamampianina, 2021).

Predictors for ch. 9**FIGURE 6** Example of spin-up for NOAA 19, AMSU-A channel 9. The predictors 1–5 correspond to predictors 0, 1, 8, 9, 10 in Table 3. [Colour figure can be viewed at wileyonlinelibrary.com]

Before production a spin-up of the VARBC coefficients was run. First a general spin-up of all coefficients for all instruments and channels and thereafter all observations were included for 10 months before the data were used actively in the assimilation. An example of a spin-up period from another experiment is shown in Figure 6. The example shows spin-up of VARBC coefficients for the predictors used for AMSU-A, channel 9 on the NOAA-19 satellite. In this case the spin-up period was fairly short and the VARBC coefficients are almost stable after 10 cycles. However, the length of the period until stabilization can vary so the choice for CERRA with a one-month spin-up period is on the safe side.

The ZTD from ground-based GNSS stations is known to have a bias that needs to be corrected for. The bias is station-dependent as well as producer-dependent due to the different software being used by different producers. The reprocessed dataset used in CERRA, however, is produced using the same software for all stations (Pacione et al., 2017).

Only predictor 0 (constant) was used for ground-based GNSS. Attempts with more predictors have been made in other studies but the results did not improve by adding more predictors (e.g., Lindskog et al., 2017).

4 | PRODUCT EVALUATION

4.1 | Climate signal

To verify that long-term trends as well as climate variability are well represented in CERRA, we compare

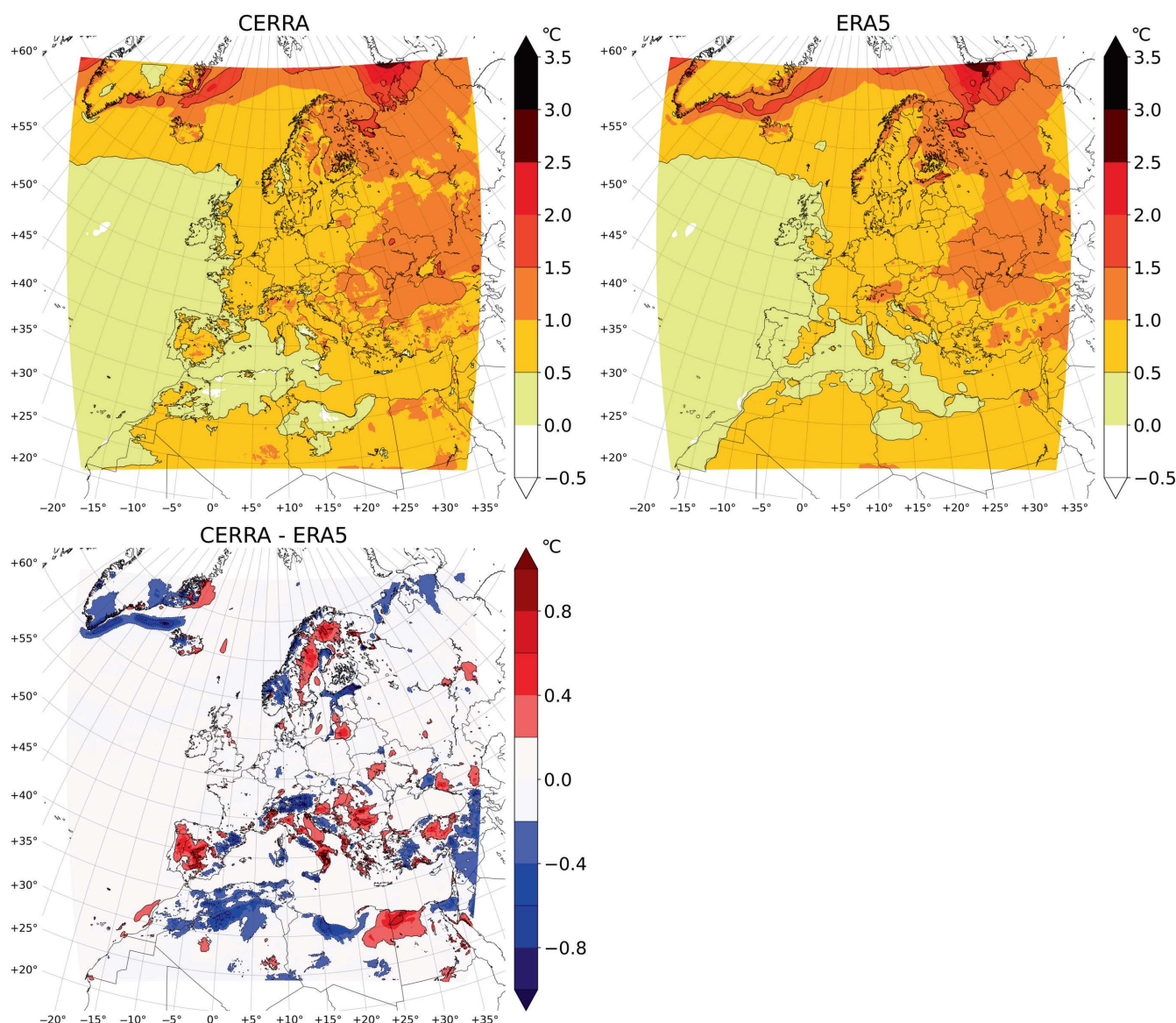


FIGURE 7 Climate change 2-m temperature signal. Differences are shown for 2003–2020 minus 1985–2002 for CERRA (top left) and ERA5 (top right). ERA5 data were interpolated to the CERRA grid for better compatibility. The bottom figure shows difference between the climate change signals in CERRA and ERA5. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/terms-and-conditions)]

long-term mean differences (2003–2020 minus 1985–2002) and annual variability of CERRA with ERA5.

For the long-term differences, the period 2003–2020 is compared to the period 1985–2022. For temperature, the climate change signal has the same pattern in CERRA and ERA5 as shown in Figure 7, where the left panel shows CERRA and the right panel ERA5. In general, the patterns and the amplitude agree very well between CERRA and ERA5. For both reanalyses, the temperature increase is more pronounced to the northeast of Europe. Some important differences in the signals are present though, for example over the Iberian Peninsula and areas with steep orography like the Alps, where the signal is stronger in CERRA. These differences are most likely attributed to

the higher resolution in CERRA. The differences over the Baltic Sea are due to the fact that CERRA uses a locally produced reanalysis dataset in the Baltic Sea area (see Section 2.3.4).

Examples of annual mean temperature series are given in Figure 8 for the cities Stockholm (Sweden), Minsk (Belarus), Paris (France) and Ajaccio (France, on Corsica). Realizations of CERRA and ERA5 are compared with observations for the locations. In general, the year-to-year variability is very well represented in both reanalyses. Cold years, such as 2010 in Stockholm, as well as warm years, for example 2011 in Paris, are well depicted in both reanalyses and the amplitudes are comparable to observations. Figure 8 shows that ERA5 has a slight cold bias

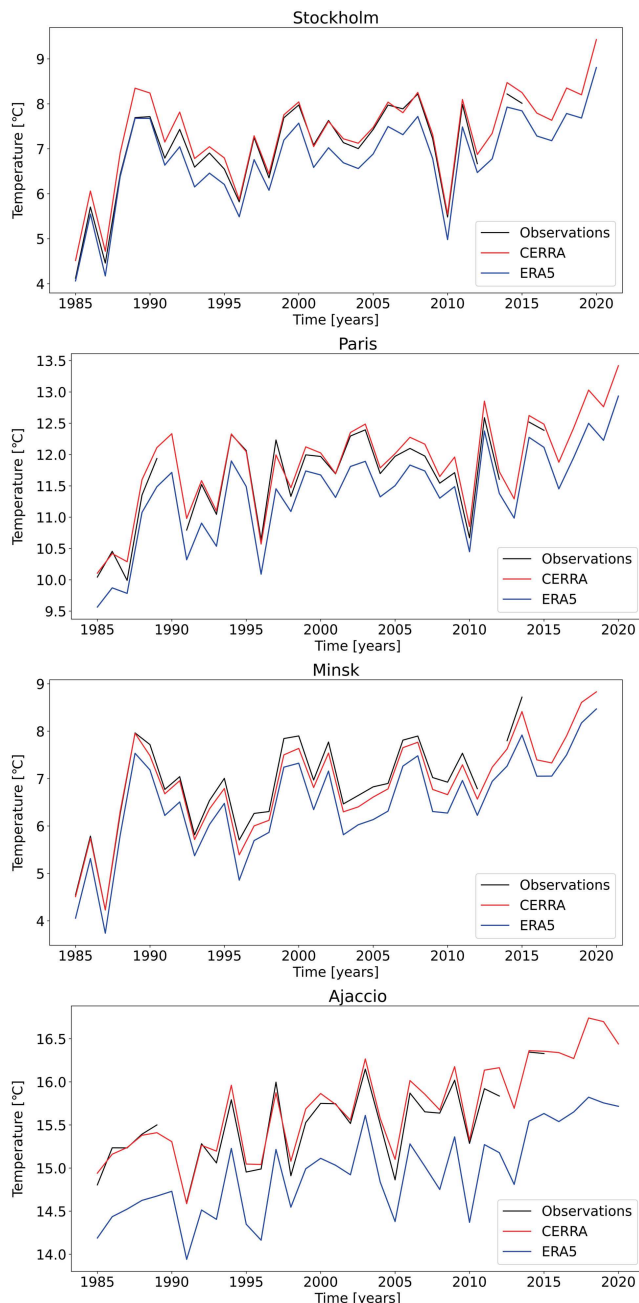


FIGURE 8 Annual mean surface temperature (T2m) for Stockholm (Sweden), Paris (France), Minsk (Belarus) and Ajaccio (France, on Corsica) computed from observations, CERRA and ERA5. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/terms-and-conditions)]

for these locations and that CERRA is generally closer to observations at these stations. In Section 4.2.1, we draw a similar conclusion for the Alpine region but it is not necessarily the case for all places in the CERRA domain.

It is also worth noting that the temperature trends for the four locations (Figure 8) over the period 1985–2021 are comparable to each other for CERRA, ERA5 and the

observations, which is in line with the comparison of the climate change signal in Figure 7.

Overall, these results indicate that CERRA has reliable temperature climate trends as well as the year-to-year variability shows good performance in CERRA.

Based on monthly mean temperatures for the four locations, Table 4 summarizes some basic statistics. The cold bias of ERA5 visible in Figure 8 is backed up with numbers, for example, ranging from -0.32 K for Stockholm to -0.66 K for Ajaccio. The bias for CERRA is much lower for all four stations. The correlation with observed monthly mean temperatures is very high for both reanalyses. This is at least to some degree due to the seasonal cycle. However, CERRA has a slightly better correlation with observations than ERA5. CERRA has smaller RMSEs as well though the RMSEs become quite similar if the biases are removed before the computation of the RMSE (Table 4).

4.2 | Seasonal verification

To verify the forecasts the HARMONIE verification tool (Monitor) has been used. With this tool a set of observations is used to do a comparison to one or several model runs, in this case CERRA and ERA5 for comparison. Model data are interpolated to the observation locations from which several standard statistical measures can be produced. Here we have chosen to display standard deviation and bias to represent the random and the systematic errors of the model as compared to observations.

Each year in the reanalysis period has been divided into four seasons; winter (December to February), spring (March to May), summer (June to August) and autumn (September to November). The verification statistics are computed for each season separately, thus each data point represents verification of a three-month period.

Many parameters could be presented but in this Section we restrict them to the surface parameters temperature, wind speed and relative humidity as examples. These variables are included in the model control vector, together with surface pressure, and can therefore be representative for the model's performance. Temperature and wind speed at 500 hPa are also shown as examples of verification at higher levels.

The forecast length presented is the three-hour forecasts. The reason is that this is the first guess for the next analysis and needs to represent the true state well. The analysis can be made to look very good if much weight is given to the observations but the system will be unbalanced and behave strange when the forecast runs. A good three-hour forecast is therefore also an indication that the system is well balanced.

	Bias		Correlation		RMSE	
	CERRA	ERA5	CERRA	ERA5	CERRA	ERA5
Stockholm	0.1	−0.32	0.98	0.98	1.34	1.37
Paris	0.09	−0.39	1.0	0.98	0.27	0.48
Minsk	−0.15	−0.54	0.98	0.98	1.75	1.81
Ajaccio	0.11	−0.66	1.0	1.0	0.36	0.88

Note: The statistics include the average bias, the correlation and the RMSE.

TABLE 4 Statistics based on monthly temperature means for CERRA and ERA5 in comparison to observations.

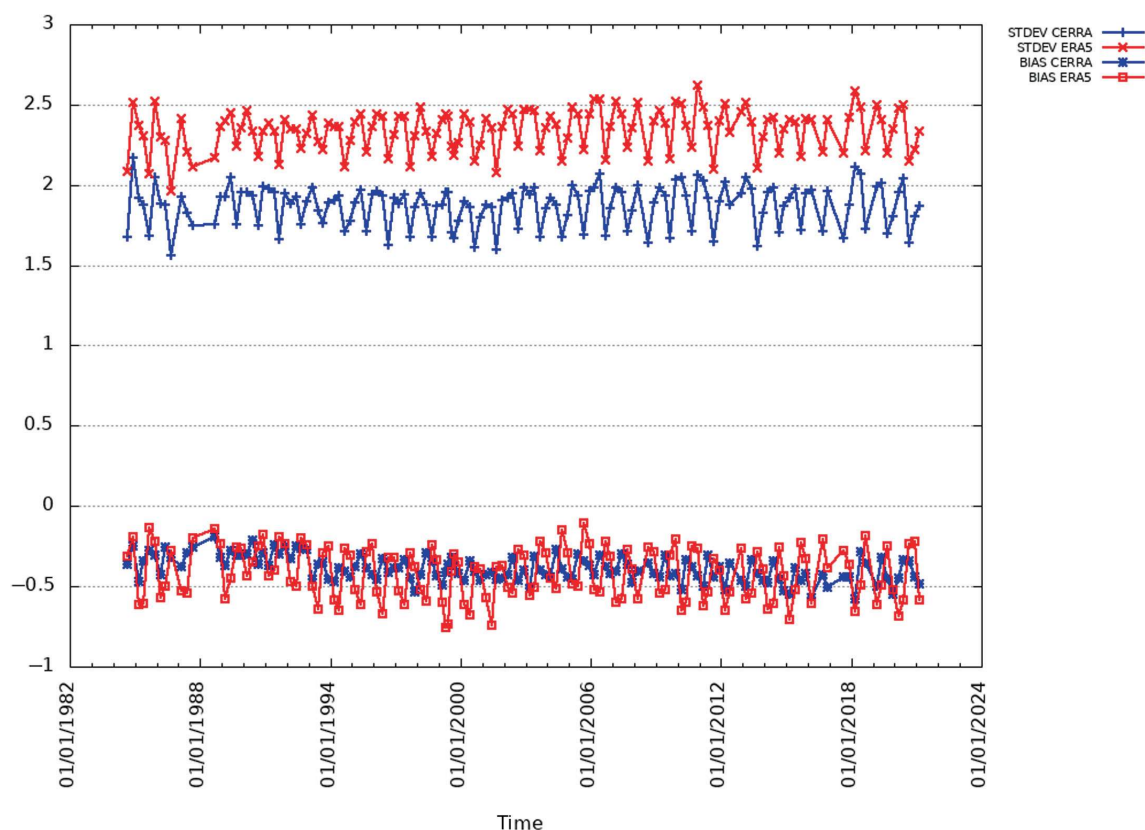


FIGURE 9 Verification of surface temperature (T2m) for the entire domain. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/qj.4764)]

4.2.1 | Surface temperature

The surface temperature (T2m) is measured two meters above the surface. Since CERRA has higher horizontal resolution and assimilates more surface observations than ERA5 it is expected that CERRA add values over ERA5 especially when it comes to small-scale features. This can also be seen in Figures 9 and 10 which show verification of three-hour forecasts over the entire domain (Figure 9), with around 1600 stations, and a smaller area over the Alps (Figure 10) with about 90 stations between the latitudes 42.8°–48.0° N and longitudes 5.35° to 16.0° E. In both cases CERRA shows a smaller standard deviation for the entire reanalysis period while the biases are about the same for the two. The bias for CERRA however, has much less seasonal dependence compared to ERA5 that

has a wider spread. For the area over the Alps (Figure 10) the added value from CERRA is larger, which is expected. Here, ERA5 shows a trend of getting closer to the observations toward the end of the period while CERRA is more or less constant in time.

Twelve-hour forecasts (not shown) indicate the same behavior with the exception that there is a slightly larger seasonal variation in CERRA bias, which is thus more similar to the ERA5 bias.

4.2.2 | Surface wind speed

The same figures as for temperature are shown for wind speed at a height of 10 m, with verification of the whole domain in Figure 11 and of the smaller area over the Alps

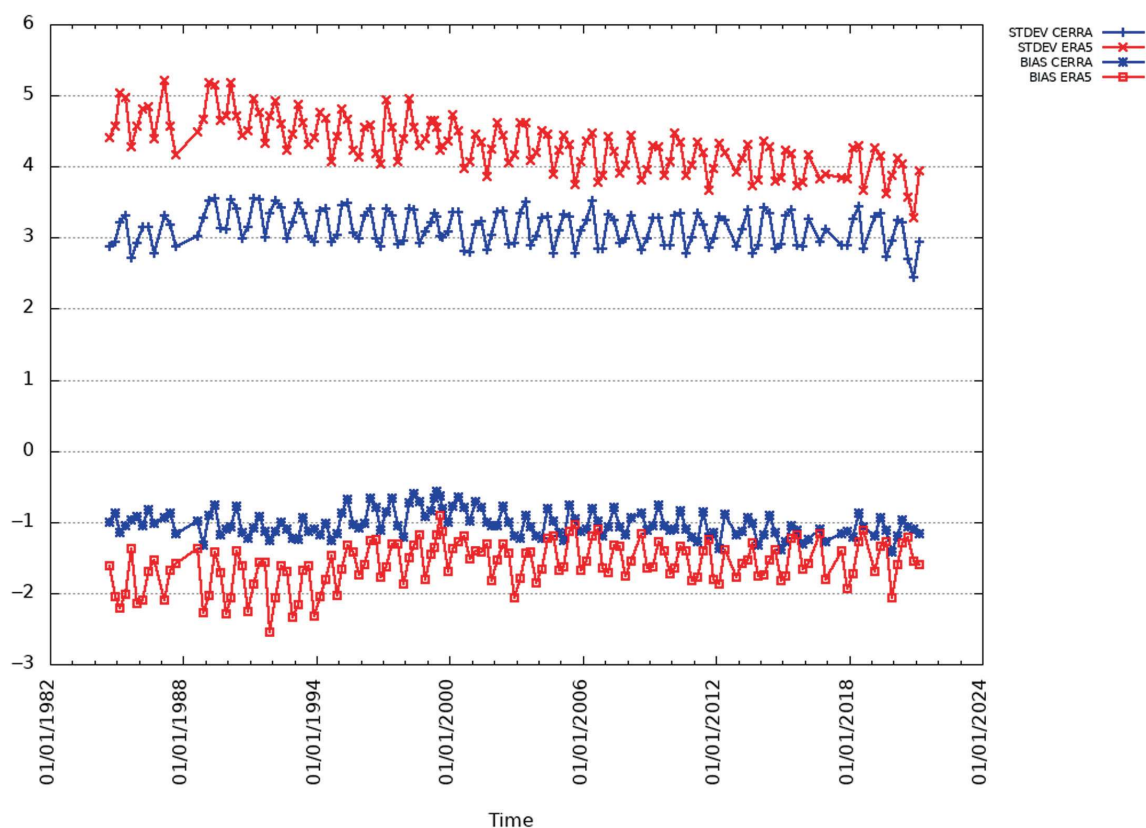


FIGURE 10 Verification of surface temperature (T2m) over the Alps (latitude 42.8°–48.0° N, longitude 5.35°–16.0° E). [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/qj.4764)]

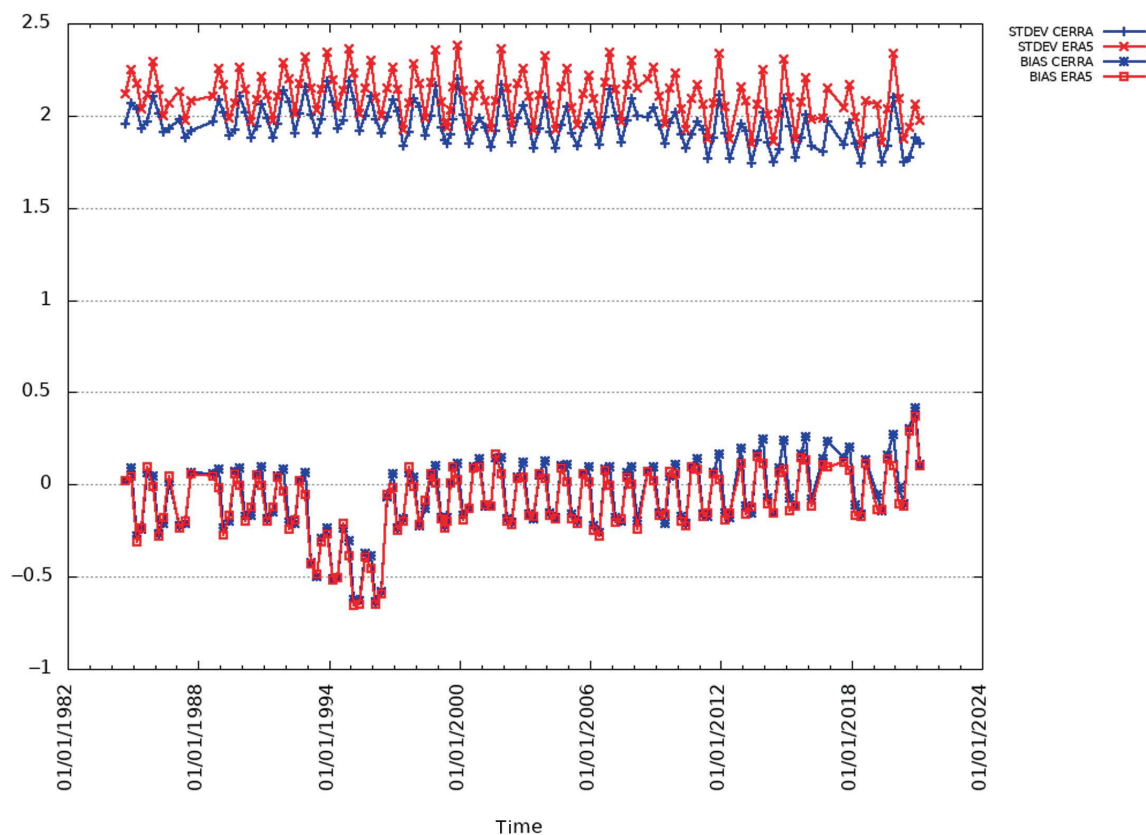


FIGURE 11 Verification of wind speed for the entire domain. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/qj.4764)]

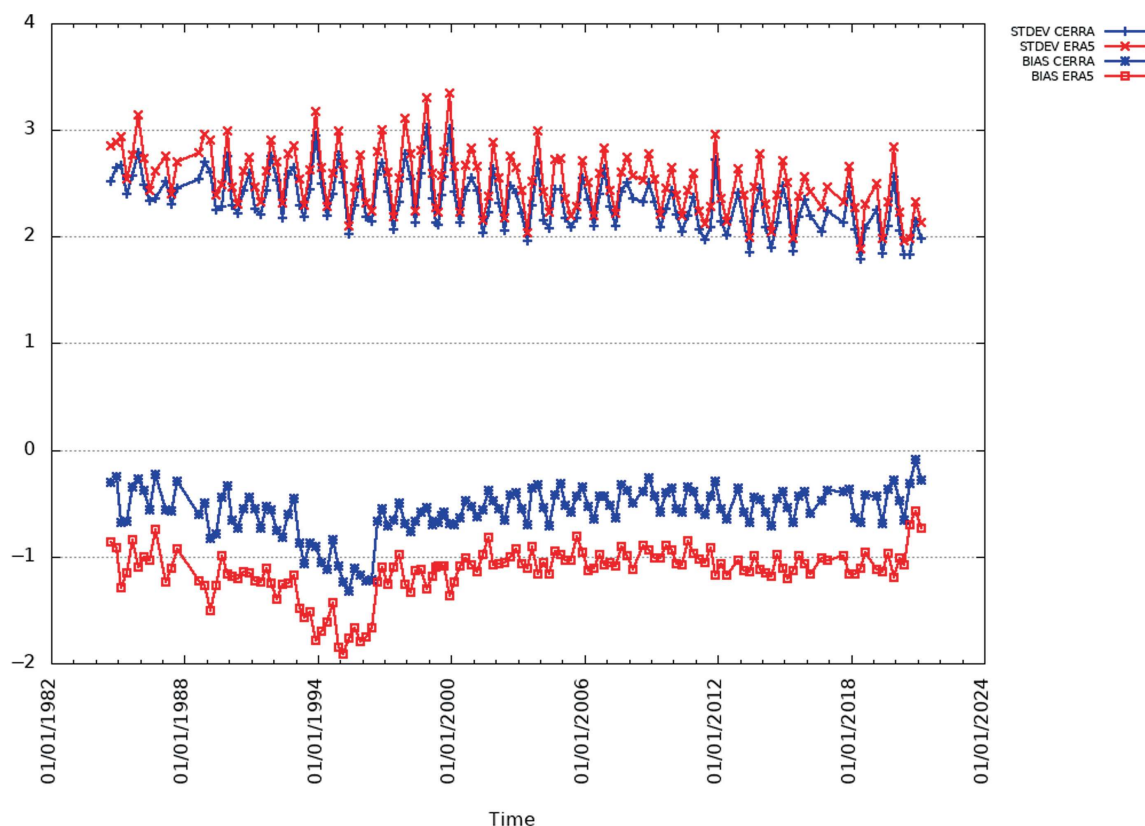


FIGURE 12 Verification of wind speed over the Alps (latitude 42.8° – 48.0° N, longitude 5.35° – 16.0° E). [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/qj.4764)]

in Figure 12. The standard deviation is better for CERRA in both cases but not as pronounced as for T2m. The bias is similar between CERRA and ERA5 when looking at the entire domain (Figure 11) while there is a much larger bias for ERA5 when looking at the Alps (Figure 12).

There is a rather strange feature in the bias verification around 1993–1996. This is probably a verification problem since it is the same for both CERRA and ERA5. This will be further investigated but it will not change any results and conclusions in this paper.

4.2.3 | Surface relative humidity

For relative humidity (RH) the results are not so good for CERRA when looking at three-hour forecasts over the whole domain as presented in Figure 13. Both standard deviation and bias indicate a slightly lower quality of CERRA compared to ERA5.

However, RH is a small-scale and rather fast-changing parameter and the large-scale verification may be a bit misleading. Verifying over a smaller area with complex terrain, again over the Alps (Figure 14), shows somewhat better results. In this area the standard deviation is similar for CERRA and ERA5 while the bias is slightly better for

CERRA even though the annual variability is rather large. For 12-hour forecasts (not shown) the standard deviation is actually better for CERRA compared to ERA5 even if the difference is rather small.

4.2.4 | The free atmosphere

Verification against radiosondes gives an indication for how the models behave in the upper levels of the atmosphere. Figure 15 shows the verification of temperature at 500 hPa for both CERRA and ERA5. The left panel shows the analysis, while the right panel shows the 12-hour forecasts. For the higher levels, verification of three-hour forecasts is not available since verification was only done two times per day to match the availability of the radiosondes. They are only available at midnight and midday and CERRA only produces long forecasts at midnight and midday.

At the analysis time, short-range temperature forecasts of CERRA are better than of ERA5 but after 12-hours the quality of the two is about the same. It is generally very hard for a regional model to beat a global model at the upper levels especially if the global model is used as boundary conditions.

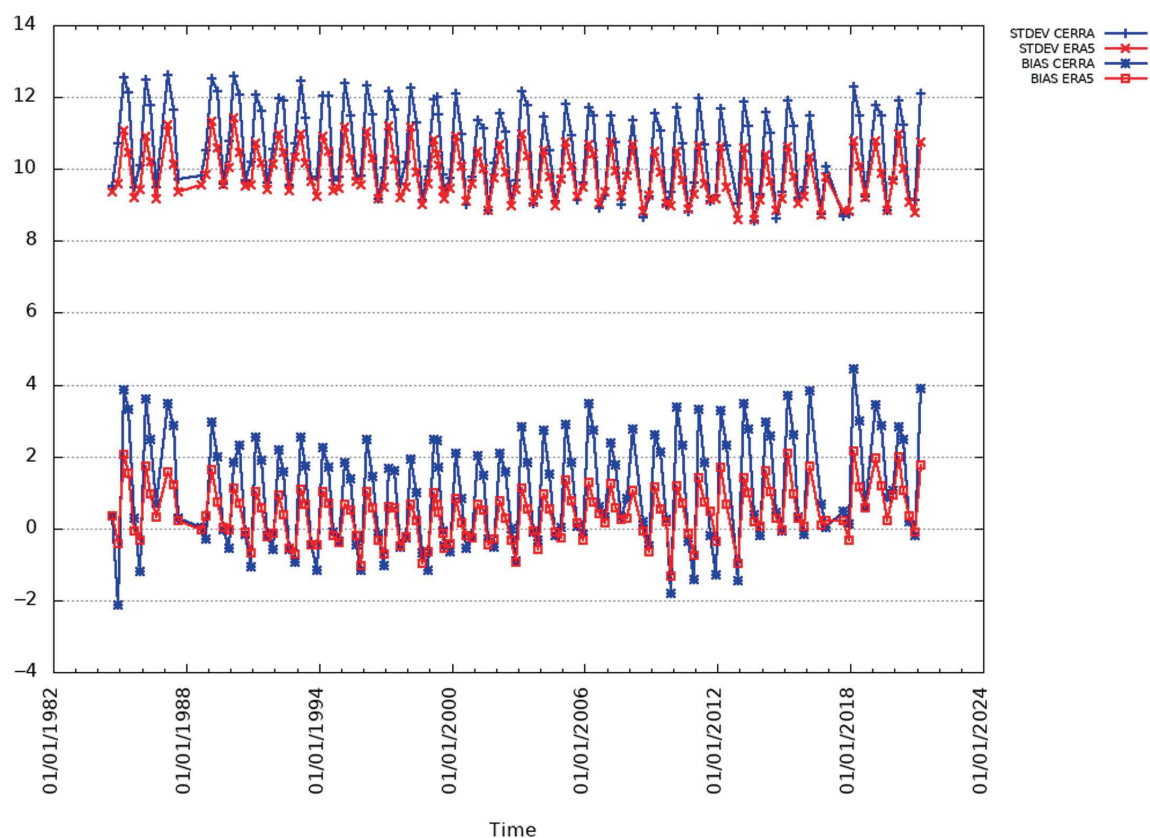


FIGURE 13 Verification of relative humidity (RH) at 2 m for the entire domain. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/terms-and-conditions)]

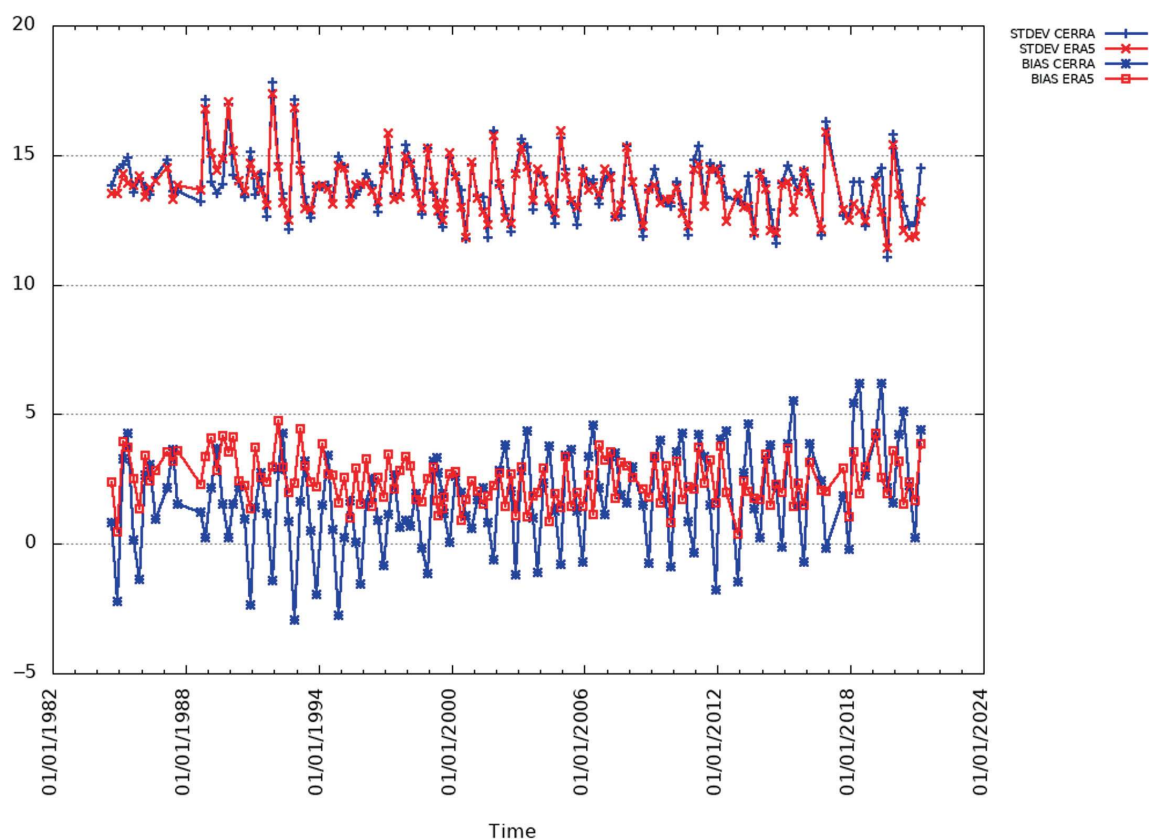


FIGURE 14 Verification of relative humidity (RH) at 2 m over the Alps (latitude 42.8°–48.0° N, longitude 5.35°–16.0° E). [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/terms-and-conditions)]

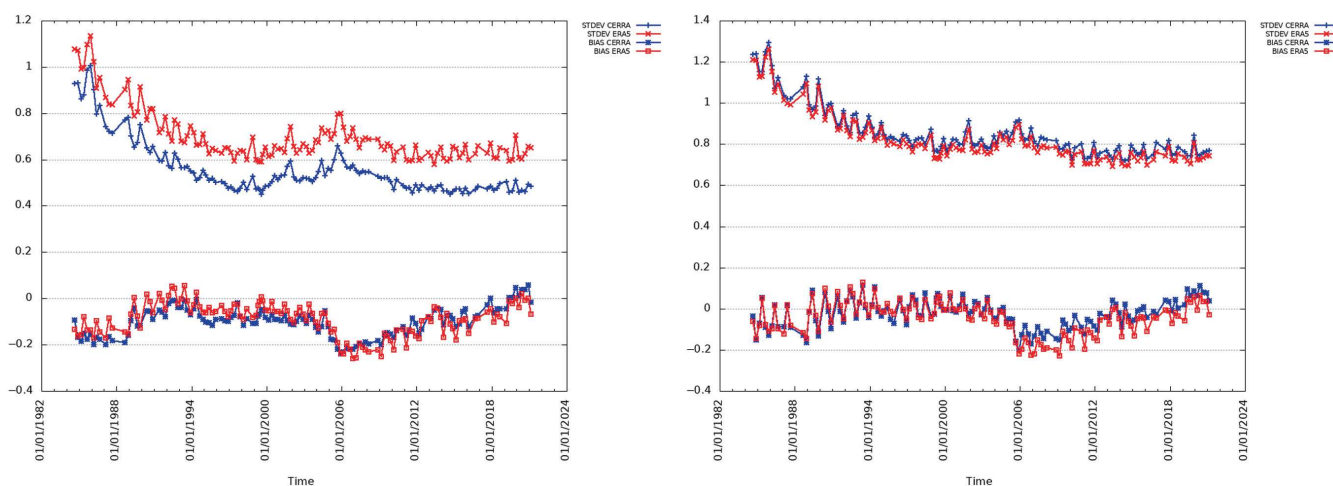


FIGURE 15 Verification of temperature at 500 hPa. Analysis in the left panel and 12-hour forecasts in the right panel. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/qj.4764)]

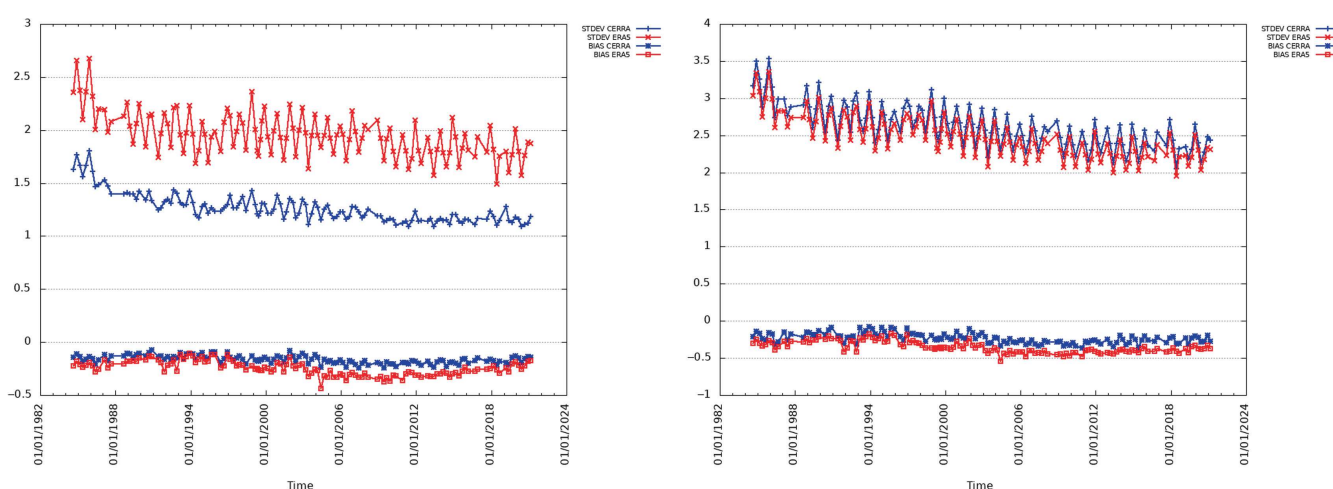


FIGURE 16 Verification of wind speed at 500 hPa. Analysis in the left panel and 12-hour forecasts in the right panel. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/qj.4764)]

The same thing can be noted for wind speed at 500 hPa presented in Figure 16. For the analysis and very short forecasts CERRA performs better than ERA5 but at 12-hour forecasts they are very similar.

Interesting to note in both Figures 15 and 16 is that the standard deviation shows a decreasing trend during the reanalysis time, which indicates an increase of quality. The reason for this is that more and more satellite observations, including GNSS-RO, become available with time that give information on upper levels of the atmosphere.

4.3 | Case study

As an example of a case study one of the most severe storms ever registered for southern Sweden was chosen. The storm, given the name Gudrun, occurred in January 2005.

The 10-m wind speed at the culmination of the storm is shown in Figure 17. A comparison between CERRA (lower panel) and ERA5 (upper panel) shows that ERA5 has in general lower wind speeds. Due to the higher spatial resolution, CERRA also captures small-scale features such as the higher wind speeds over the big lakes and reduced wind speed over islands.

Compared to observations CERRA wind speed is generally closer to the observed values as shown in Figure 18. Figure 18 shows time series (analysis plus one- and two-hour forecasts) for stations Hanö and Måseskär (indicated as locations A and B in Figure 17) for CERRA, ERA5 as well as the observed wind speed. For Hanö CERRA wind speed is closer to the observations for the whole period compared to ERA5. For Måseskär the two reanalyses are rather similar except for the peak wind speed where ERA5 is too low but CERRA

2005-01-08 18UTC

ERA5

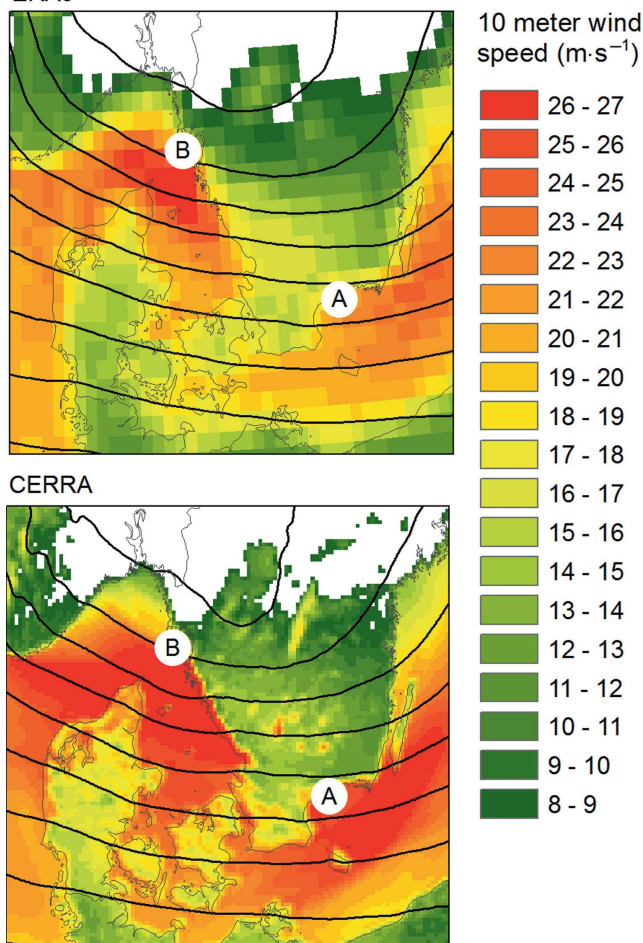


FIGURE 17 10-m wind at 1800 UTC on January 8, 2005 for both ERA5 and CERRA. Time series of locations A and B are shown in Figure 18. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/qj.4764)]

is close to the observed value. Included for comparison is also wind speed from UERRA-HARMONIE (analysis plus 1–5-hour forecasts). The reason UERRA-HARMONIE is “spikey” is because there was an error in the turbulent kinetic energy initialization. Due to this problem, users are recommended not to use the shortest forecast (1–3 hours) of UERRA-HARMONIE winds.

4.4 | Seasonal variability

The seasonal variability of CERRA, ERA5 and UERRA-HARMONIE was investigated over Sweden for the time period 1989–2018. The verification was made for the six-hour forecasts against all available Swedish meteorological observations. For the 2-m temperature and 10-m wind speed we find that CERRA adds clear value to UERRA-HARMONIE and ERA5, the latter as

also shown in Section 4.2, with improved statistical scores, demonstrating the benefits of the sophisticated and higher resolution in the regional reanalysis. For the RH on the other hand CERRA is clearly better than UERRA-HARMONIE but not better than ERA5 as also shown above.

In Figure 19 the RMSE is shown for the three parameters and model runs and it is clear that the variability is larger during winter for temperature and wind speed. RH on the other hand shows the largest variability during spring.

4.5 | Precipitation

During a period over 10 years (1990–2000), precipitation from a network of daily independent rain gauge stations in France was compared to the reanalyses ERA5, CERRA, CERRA-Land and UERRA-Land (also known as UERRA MESCAN/SURFEX). The number of observation was around 2000 stations/day. From this dataset different verification scores were calculated such as the Heidke Skill Score (HSS) and the false alarm rate (FAR). The HSS gives information on analysis accuracy with a perfect score of 1 while the FAR gives the probability that an expected event did not occur. FAR is between 0 and 1 where 0 is a perfect forecast.

In Figure 20 the scores are presented and they show CERRA is better than ERA5 for all precipitation intensities. It also shows that precipitation analysis from the two surface analyses CERRA-Land and UERRA-Land has clearly better scores, with CERRA-Land giving the best scores. It improves the representation of precipitation over France. Furthermore, the daily surface precipitation of CERRA has a smaller RMSE than that from ERA5 (4.09 mm vs 4.32 mm). This added value has contributed to improve the precipitation analysis in CERRA-Land, which is a combination of CERRA and precipitation observations, and therefore also in the calculation of snow depth.

5 | TECHNICAL IMPLEMENTATION

The production of a reanalysis needs extensive computational resources. For CERRA-EDA and CERRA, the High Performance Computing Facility (HPCF) of ECMWF in Reading was used. Even if this is a powerful system it would take too much time to produce the entire series in one go. The production was therefore split into several streams. In total, CERRA-EDA was produced

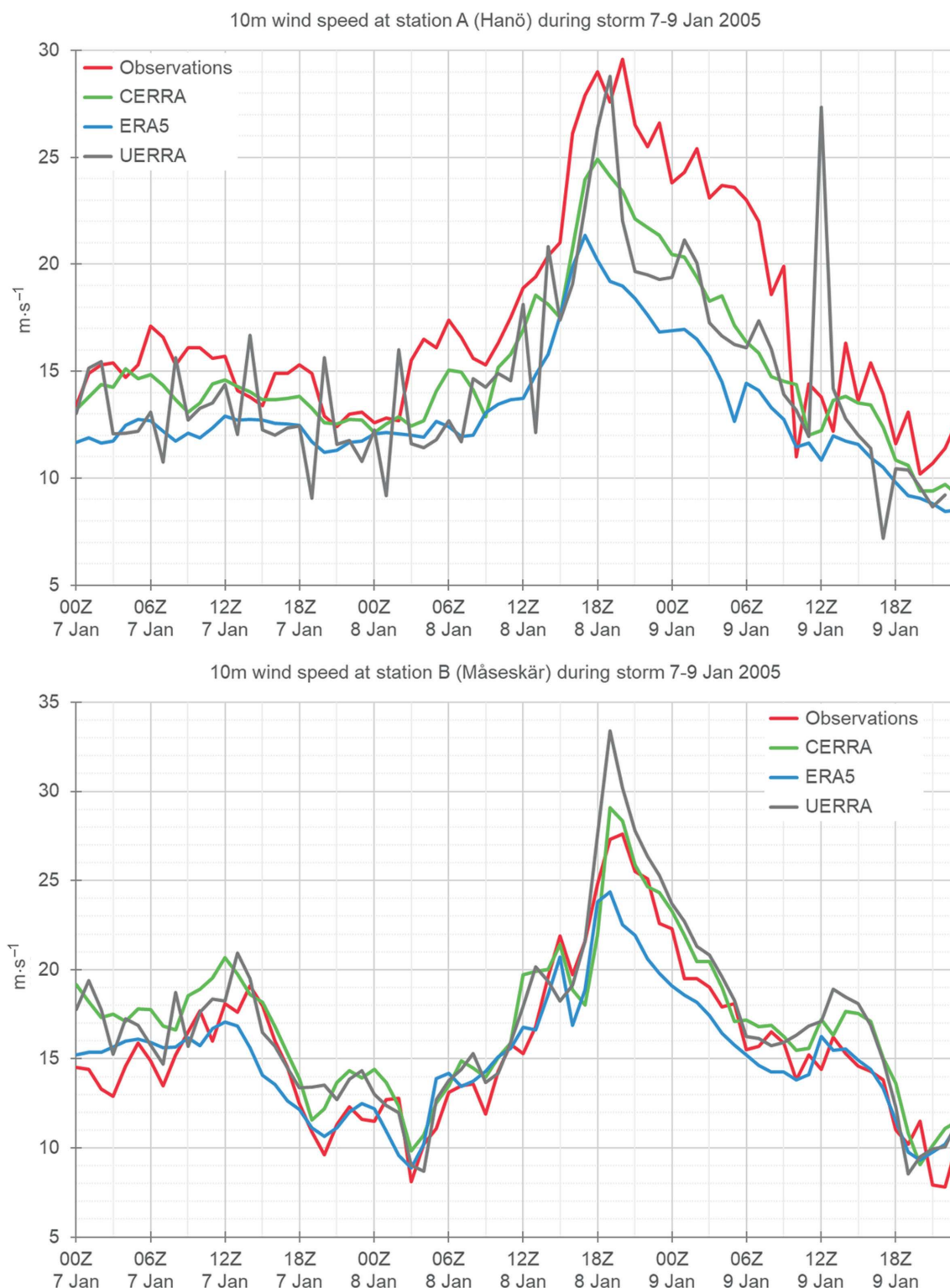


FIGURE 18 Time series of 10-m wind speed during the storm Gudrun. The locations of the two stations are indicated in Figure 17. [Colour figure can be viewed at wileyonlinelibrary.com]

with 11 streams and CERRA with 13 streams. In addition, each stream started with a spin-up of one year to get the model state into balance. This is illustrated in Figure 21.

With the production split up into so many streams it took more than 1.5 years (March 2020 to September 2021) to complete the reanalysis. The average production speed (referred to as throughput) was heavily dependent

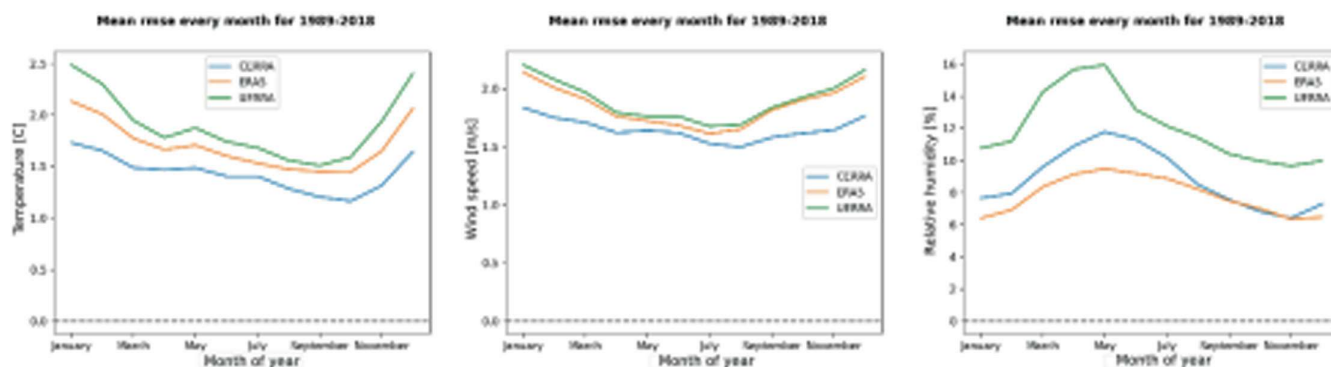


FIGURE 19 Monthly root-mean-squared error for 2-m temperature (left panel), 10-m wind speed (middle panel) and 2-m relative humidity (right panel) for the time period 1989 to 2018 over Sweden. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/qj.4764)]

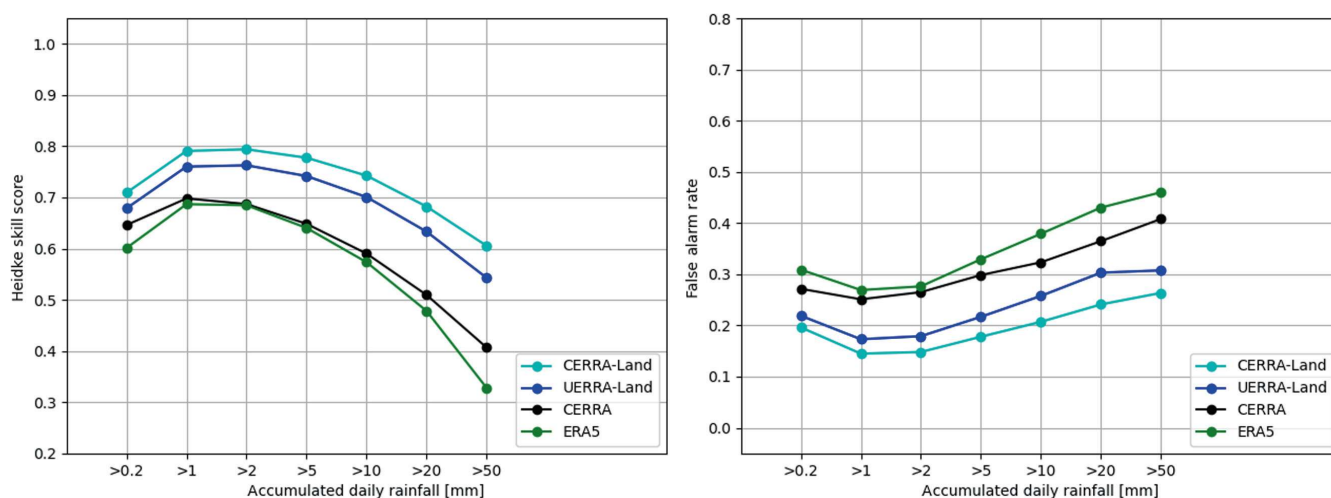


FIGURE 20 Heidke Skill Score (HSS, left panel) and false alarm rate (FAR, right panel) for CERRA-Land precipitation analysis, UERRA-Land precipitation analysis, CERRA precipitation and ERA5 precipitation. For HSS the perfect score is 1 whereas for FAR it is 0. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/qj.4764)]

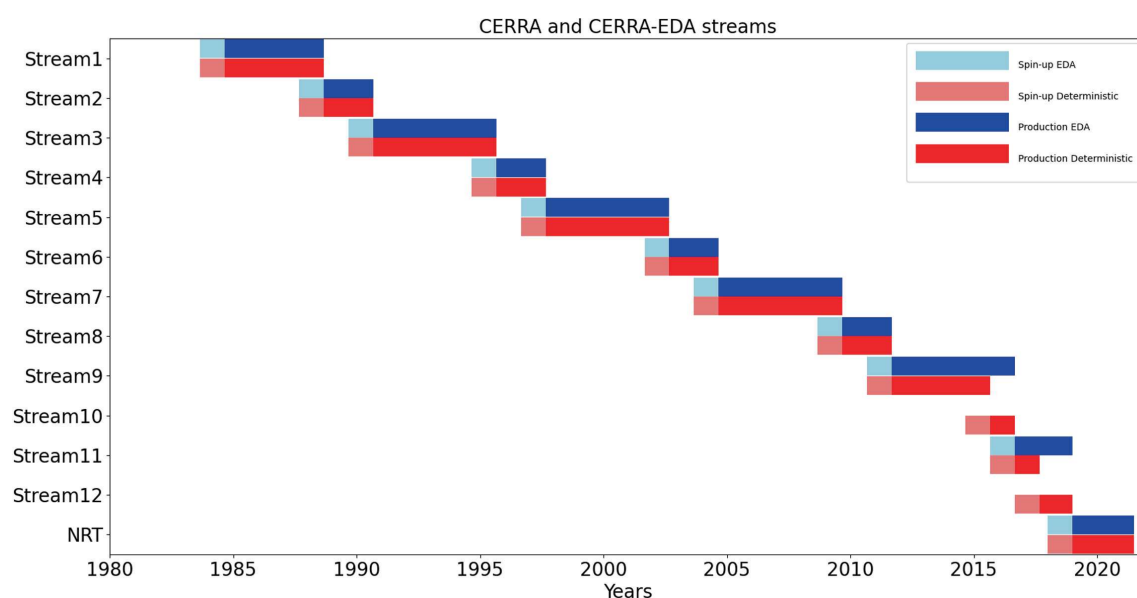


FIGURE 21 CERRA and CERRA-EDA (ensemble data assimilation) streams including the corresponding spin-up periods. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/qj.4764)]

on the general load of the HPCF as well as on the number of assimilated observations. With less observations in the early period, the throughput of the early streams was generally somewhat higher than for the later periods.

In the HARMONIE-ALADIN system, the time critical parts are the creation of the analysis and the running of the forecasts. The preparation of input data as well as the post-processing was able to run much faster. Whereas the workload of the input data preparation was simply much smaller, the advantage of the post-processing was that all cycles are independent and can therefore run in parallel. That is only partly the case for the analyses and forecasts, where each cycle is depending on input from the previous cycle.

6 | CONCLUSIONS

A 37-year-long regional reanalysis has been produced using a newly developed model system. The new system combines different versions and physics packages of the HARMONIE system used operationally in several European countries. However, there is no operational HARMONIE system running for the CERRA domain. Operational systems are global or covering a much smaller area for regional setups.

In CERRA a new way of constructing the **B** matrix has been used with good results. The ensemble was used in the calculation of the **B** matrix which is updated every two days. This has not been used by the HARMONIE users previously and investigations are ongoing to see if the same method can be used in operational runs.

Boundaries as well as observations data have been taken from the global reanalysis ERA5. However, some observation types not used globally, like GBGNSS, were used in CERRA and in some cases different datasets, for example, reprocessed datasets, have been used. This leads to better signal-to-noise ratio and thereby more observations included. A large effort has also been made to collect local observations not stored in the global databases. This was done in collaboration with the parallel Arctic regional reanalysis CARRA.

Results presented here show added value for most parameters in most cases. It should be noted that when areas or time scales are aggregated, it is not expected that CERRA outperforms the global ERA5 reanalysis for all parameters and levels. The advantage of CERRA comes with the higher resolution and therefore added value is expected for small regions and on shorter time scales. As shown in Section 4.2, the added value of CERRA is more clear when selecting a smaller area with complex terrain, such as the Alps. The results shown are seasonal averages

so an even higher degree of added value can be expected for monthly or hourly verification. As also shown, in the upper air fields, the advantage of CERRA is diluted in the longer forecasts at higher levels.

CERRA data are available for users through the Copernicus CDS (Schimanke et al., 2021). From the CDS a comprehensive user guide can also be accessed. The user guide includes also a section on known issues.

For most surface parameters, for example, precipitation, it is recommended to use the output from CERRA-Land. In fact some parameters, like soil moisture and temperature, are not available from CERRA in the CDS. The two-dimensional CERRA-Land analysis is run with more advanced surface parameterizations as well as additional observations, especially of precipitation.

For the next generation of CERRA there are a few lessons learned that should be considered. One example is that development always takes longer than expected. In the case of CERRA, stepping away from operational settings made the development a bit slower than anticipated due to extra tuning of some parameters. Another important lesson is that running 37 years in 13 parallel hi-resolution streams and 12 EDA streams, each running with 10 members, puts a lot of pressure on the computer resources and disk usage. A lot of time was spent by the CERRA team together with systems people at ECMWF on optimizing the system after the production was started. Much of this could have been done in advance to be better prepared.

Plans are to re-start the production of CERRA and CERRA-EDA to catch up with the current date and continue with timely updates for a few more years. In addition, a back extension is planned in order to cover the same time period as UERRA-HARMONIE, that is, starting from 1961.

ACKNOWLEDGEMENTS

The authors would like to acknowledge all the staff at SMHI, Météo-France and MET-Norway that somehow were involved in the development and production of the CERRA dataset. It was a large project that touched upon many people. Monica Wallgren deserves a special mention for excellent handling of all economical issues. The CARRA team should be acknowledged for a very constructive and inspiring collaboration. We would also especially like to thank Bojan Kasic and Dominique Lucas at ECMWF for help with the optimizing of the system as well as Richard Mladek and Sebastien Villaume for help with the archiving. Finally, a very big thanks to Cornel Soci for being our contact point for ECMWF and Copernicus. This work has been funded by the Copernicus Climate Change Service through Copernicus Europe Grant

Number C3S_322Lot1. ECMWF implements this Service on behalf of the European Commission.

DATA AVAILABILITY STATEMENT

Data is openly available through the Copernicus Climate Data Store.

ORCID

Martin Ridal  <https://orcid.org/0000-0002-2716-1571>

REFERENCES

- Auligné, T., McNally, A.P. & Dee, D.P. (2007) Adaptive bias correction for satellite data in a numerical weather prediction system. *Quarterly Journal of the Royal Meteorological Society*, 133, 631–642.
- Bauer, P., Radnoti, G., Healy, S. & Cardinali, C. (2014) GNSS radio occultation constellation observing system experiments. *Monthly Weather Review*, 142, 555–572.
- Bazile, E., Marquet, P., Bouteloup, Y. & Bouyssel, F. (2011) The turbulent kinetic energy (TKE) scheme in the NWP models at Météo-France. In: *Workshop on diurnal cycles and the stable boundary layer*. Reading, UK: ECMWF, pp. 127–136.
- Bechtold, P., Bazile, E., Guichard, F., Mascart, P. & Richard, E. (2001) A mass-flux convection scheme for regional and global models. *Quarterly Journal of the Royal Meteorological Society*, 127(573), 869–886.
- Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., de Rooy, W. et al. (2017) The Harmonie-AROME model configuration in the ALADIN-HIRLAM NWP system. *Monthly Weather Review*, 145, 1919–1935. Available from: <https://doi.org/10.1175/MWR-D-16-0417.1>
- Berre, L. (2000) Estimation of synoptic and mesoscale forecast error covariances in a limited-area model. *Monthly Weather Review*, 128, 644–667.
- Bollmeyer, C., Keller, J.D., Ohlwein, C., Wahl, S., Crewell, S., Friederichs, P., Hense, A., Keune, J., Kneifel, S., Pscheidt, I., Redl, S. & Steinke, S. (2015) Towards a high-resolution regional reanalysis for the European CORDEX domain. *Q. J. R. Meteorol. Soc.*, 141, 1–15. <https://doi.org/10.1002/qj.2486>
- Bougeault, P. (1985) A simple parameterization of the large-scale effects of cumulus convection. *Monthly Weather Review*, 113, 2108–2121.
- Bougeault, P. & Lacarrere, P. (1989) Parameterization of orography-induced turbulence in a Mesobeta-scale model. *Monthly Weather Review*, 117, 1872–1890.
- Bouteloup, Y., Bouyssel, F. & Marquet, P. (2005) Improvements of Lopez's prognostic large scale cloud and precipitation scheme. *ALADIN Newsletter*, 28, 66–73.
- Bromwich, D., Bai, L., Hines, K., Wang, S., Liu, Z., Lin, H. et al. (2012) Arctic system reanalysis (ASR) project research data archive at the national center for atmospheric research, computational and information systems laboratory. <https://doi.org/10.5065/D6K072B5>
- Bromwich, D., Wilson, A.B., Bai, L., Liu, Z., Barlage, M., Shih, C.-F. et al. (2017) The Arctic system reanalysis, version 2. *Bulletin of the American Meteorological Society*, 99, 805–828. Available from: <https://doi.org/10.1175/BAMS-D-16-0215.1>
- Brousseau, P., Bouttier, F., Hello, G., Seity, Y., Berre, L., Montmerle, T. et al. (2008) A prototype convective-scale data assimilation system for operation: the AROME-RUC. *HIRLAM Technical Report*, 68, 23–30.
- Brousseau, P., Berre, L., Bouttier, F. & Desroziers, G. (2012) Flow-dependent background-error covariances for a convective-scale data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 138, 310–322.
- Catry, B., Geleyn, J.-F., Bouyssel, F., Cedilnik, J., Brožková, R., Derková, M. et al. (2008) A new sub-grid scale lift formulation in a mountain drag parameterization scheme. *Meteorologische Zeitschrift*, 17(2), 193–208. Available from: <https://doi.org/10.1127/0941-2948/2008/0272>
- Cuxart, J., Bougeault, P. & Redelsperger, J.-L. (2000) A turbulence scheme allowing for mesoscale and large-eddy simulations. *Quarterly Journal of the Royal Meteorological Society*, 126, 1–30. Available from: <https://doi.org/10.1002/qj.49712656202>
- Dahlgren, P., Landelius, T., Källberg, P. & Gollvik, S. (2016) A high-resolution regional reanalysis for Europe. Part 1: three-dimensional reanalysis with the regional High-resolution limited-area model (HIRLAM). *Quarterly Journal of the Royal Meteorological Society*, 142, 2119–2131.
- Danielson, J.J. & Gesch, D.B. (2011) Global multi-resolution terrain elevation data 2010 (GMTED2010): U.S. Geological Survey Open-File Report 2011–1073. 26 p., <https://doi.org/10.3133/ofr20111073>
- Dee, D.P. (2004) *Variational bias correction of radiance data in the ECMWF system, in proceedings of workshop on assimilation of high spectral resolution sounders in NWP*, Reading, UK, 28 June–1 July 2004. Reading, UK: ECMWF, pp. 97–112.
- Dee, D.P. & co-authors. (2011) The ERA-interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597.
- Dee, D.P., Balmaseda, M., Balsamo, G., Simmons, R.E.A.J. & Thepaut, J.-N. (2014) Towards a consistent reanalysis of the climate system. *Bulletin of the American Meteorological Society*, 95, 1235–1248.
- Derber, J.C. & Wu, W.-S. (1998) The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Monthly Weather Review*, 126, 2287–2299.
- Donlon, C.J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E. & Wimmer, W. (2012) The Operational Sea surface temperature and sea ice analysis (OSTIA) system. *Remote Sensing of the Environment*, 116, 140–158. Available from: <https://doi.org/10.1016/j.rse.2010.10.017>
- El-Said, A., Brousseau, P., Ridal, M. & Randriamampianina, R. (2022) Towards a full flow-dependence: new temporally varying EDA quotient functionality to estimate background errors in CERRA. *Journal of Advances in Modeling Earth Systems*, 14, 1–27. Available from: <https://doi.org/10.1029/2021MS002637>
- Faroux, S., Kaptué Tchuenté, A.T., Roujean, J.-L., Masson, V., Martin, E. & Le Moigne, P. (2013) ECOCLIMAP-II/Europe: a twofold database of ecosystems and surface parameters at 1 km resolution based on satellite information for use in land surface, meteorological and climate models. *Geoscientific Model Development*, 6, 563–582. Available from: <https://doi.org/10.5194/gmd-6-563-2013>
- Fischer, C., Montmerle, T., Berre, L., Auger, L. & Ștefănescu, S.E. (2005) An overview of the variational assimilation in the ALADIN/France numerical weather-prediction system. *Quarterly Journal of the Royal Meteorological Society*, 131, 3477–3492.

- Fouquart, Y. & Bonnel, B. (1980) Computations of solar heating of the earth's atmosphere—a new parameterization. *Beitraege Zur Physik der Atmosphaere*, 53, 35–62.
- Gleeson, E., Whelan, E. & Hanley, J. (2017) Met Éireann high resolution reanalysis for Ireland. *Advances in Science and Research*, 14, 49–61.
- Gustafsson, N., Berre, L., Hörnquist, S., Huang, X.-Y., Lindskog, M., Navascués, B. et al. (2001) Three-dimensional variational data assimilation for a limited area model. Part I: general formulation and the background error constraint. *Tellus*, 53A, 425–446.
- Haimberger, L., Blaszek, M., Brönnimann, S., Dahlgren, P., Hersbach, H., Jourdain, S. et al. (2017) Bias adjustments for global radiosonde data, ERA-CLIM2 deliverable 4.2. Available at: <http://www.era-clim2.eu/products>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J. et al. (2020) The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049.
- Hordoir, R., Axell, L., Höglund, A., Dieterich, C., Frnsner, F., Gröger, M. et al. (2019) Nemo-Nordic 1.0: a NEMO-based ocean model for the Baltic and north seas – research and operational applications. *Geoscientific Model Development*, 12, 363–386. Available from: <https://doi.org/10.5194/gmd-12-363-2019>
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L. et al. (1996) The ncep/ncar 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 77, 437–471.
- Kobayashi, S., Ota, Y., Harada, Y., Ebata, A., Moriya, M., Onoda, H. et al. (2015) The jra-55 reanalysis: general specifications and basic characteristics. *Journal of the Meteorological Society of Japan*, 93, 5–48.
- Kourzeneva, E. (2010) External data for lake parameterization in numerical weather prediction and climate modelling. *Boreal Environment Research*, 15, 165–177.
- Koltzow, M., Schyberg, H., Støylen, E. & Yang, X. (2022) Value of the Copernicus Arctic regional reanalysis (CARRA) in representing near-surface temperature and wind speed in the north-east European Arctic. *Polar Research*, 41, 1–15. Available from: <https://doi.org/10.33265/polar.v41.8002>
- Lindskog, M., Gustafsson, N., Navascués, B., Mogensen, K.S., Huang, X.-Y., Yang, X. et al. (2001) Three-dimensional variational data assimilation for a limited area model. Part II: observation handling and assimilation experiments. *Tellus*, 53A, 447–468.
- Lindskog, M., Ridal, M., Thorsteinsson, S. & Ning, T. (2017) Data assimilation of GNSS zenith total delays from a Nordic processing centre. *Atmospheric Chemistry and Physics*, 17, 13983–13998.
- Liu, Y., Meier, H.E.M. & Axell, L. (2013) Reanalyzing temperature and salinity on decadal time scales using the ensemble optimal interpolation data assimilation method and a 3-D ocean circulation model of the Baltic Sea. *Journal of Geophysical Research, Oceans*, 118, 5536–5554. Available from: <https://doi.org/10.1002/jgrc.20384>
- Lopez, P. (2002) Implementation and validation of a new prognostic large-scale cloud and precipitation scheme for climate and data-assimilation purposes. *Quarterly Journal of the Royal Meteorological Society*, 128, 229–257.
- Marquet, P., Descamps, L. & Bouyssel, F. (2019) A new grid point storm control scheme in the ARPEGE NWP model. In: *WMO CAS/JSC, WGNE Blue Book*, Vol. 49. Geneva: WMO, pp. 4–11 http://bluebook.meteoinfo.ru/uploads/2017/docs/04_MARQUET_Pascal_GridPointStormControl.pdf
- Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P.C., Ebisuzaki, W. et al. (2006) North American regional reanalysis. *Bulletin of the American Meteorological Society*, 87, 343–360.
- Mlawer, E.J., Taubman, S.J., Brown, P.D., Iacono, M.J. & Clough, S.A. (1997) Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research*, 102, 16663–16682. Available from: <https://doi.org/10.1029/97JD00237>
- Montmerle, T. & Berre, L. (2010) Diagnosis and formulation of heterogeneous background-error covariances at the mesoscale. *Quarterly Journal of the Royal Meteorological Society*, 136, 1408–1420.
- Morcrette, J.J. (1991) Radiation and cloud radiative properties in the European Centre for Medium Range Weather Forecasts forecasting system. *Journal of Geophysical Research: Atmospheres*, 96, 9121–9132.
- Nachtergaele, F., Velthuisen, H., Verelst, L. & Wilberg, D. (2012) Harmonized World Soil Database Version 1.2.
- Pacione, R., Pace, B., Vedel, H., de Haan, S., Lanotte, R. & Vespe, F. (2017) Combination methods of tropospheric time series. *Advance in Space Research*, 47, 323–335.
- Poli, P., Healy, S.B. & Dee, D.P. (2010) Assimilation of global positioning system radio occultation data in the ECMWF ERA-interim reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 138, 1972–1990.
- Poulsen, J.W. & Berg, P. (2012) “More details on HBM—general modelling theory and survey of recent studies”. DMI Technical Report No. 12-16. <http://beta.dmi.dk/fileadmin/Rapporter/TR/tr12-16.pdf>
- Ridal, M., Körnich, H., Olsson, E. & Andrae, U. (2016) Preliminary report of the first period of the RA. UERRA deliverable D2.6. <http://www.uerra.eu>.
- Ridal, M., Olsson, E., Undén, P., Zimmermann, K. & Ohlsson, A. (2017) HARMONIE reanalysis report of results and dataset. UERRA deliverable D2.7. <http://www.uerra.eu>.
- Rienecker, M., Suarez, M., Gelaro, R., Todling, R., Bacmeister, J., Liu, E. et al. (2011) Merra: Nasas modern-era retrospective analysis for research and applications. *Journal of Climate*, 24, 3624–3648.
- Sanchez Arriola, J., Lindskog, M., Thorsteinsson, S. & Bojarova, J. (2016) Variational bias correction of GNSS ZTD in the HARMONIE modeling system. *Journal of Applied Meteorology and Climatology*, 55, 1259–1276.
- Schimanke, S., Ridal, M., Le Moigne, P., Berggren, L., Undén, P., Randriamampianina, R. et al. (2021) CERRA sub-daily regional reanalysis data for Europe on model levels from 1984 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), <https://doi.org/10.24381/cds.7c27fd20>
- Spanghel, T., Borsche, M., Niermann, D., Kaspar, F., Schimanke, S., Brienens, S. et al. (2023) Intercomparing the quality of recent reanalyses for offshore wind farm planning in Germany's exclusive economic zone of the North Sea. *Advances in Science and Research*, 20, 109–128. Available from: <https://doi.org/10.5194/asr-20-109-2023>
- Smith, R.N.B. (1990) A scheme for predicting layer clouds and their water content in a general circulation model. *Quarterly Journal of the Royal Meteorological Society*, 116, 435–460.
- Taillefer, F. (2002) CANARI—Technical Documentation—Based on ARPEGE cycle CY25T1 (AL25T1 for ALADIN). Available at <http://www.cnrm.meteo.fr/aladin/>,

- Termonia, P., Fischer, C., Bazile, E., Bouyssel, F., Brožková, R., Bénard, P. et al. (2018) The ALADIN system and its canonical model configurations AROME CY41T1 and ALARO CY40T1. *Geoscientific Model Development*, 11, 257–281.
- Uppala, S.M., Kållberg, P.W., Simmons, A.J. et al. (2005) The era-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, 131, 2961–3012.
- Valkonen, T., Schyberg, H. & Figa-Saldana, J. (2017) Assimilating advanced Scatterometer winds in a high-resolution limited area model over northern Europe. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10, 2394–2405. Available from: <https://doi.org/10.1109/JSTARS.2016.2602889>
- Vega, M.L., Fransson, T. & Kullberg, C. (2021) The effects of four decades of climate change on the breeding ecology of an avian sentinel species across a 1500-km latitudinal gradient are stronger at high latitudes. *Ecology and Evolution*, 11, 6233–6247. Available from: <https://doi.org/10.1002/ece3.7459>
- Verrelle, A., Glinton, M., Bazile, E., Le Moigne, P., Randriamampianina, R., Ridal, M. et al. (2022) CERRA-Land sub-daily regional reanalysis data for Europe from 1984 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). <https://doi.org/10.24381/cds.a7f3cd0b>
- Wang, Z.Q. & Randriamampianina, R. (2021) The impact of assimilating satellite radiance observations in the Copernicus European regional reanalysis (CERRA). *Remote Sensing*, 13, 1–28. Available from: <https://doi.org/10.3390/rs13030426>

How to cite this article: Ridal, M., Bazile, E., Le Moigne, P., Randriamampianina, R., Schimanke, S., Andrae, U. et al. (2024) CERRA, the Copernicus European Regional Reanalysis system. *Quarterly Journal of the Royal Meteorological Society*, 1–27. Available from: <https://doi.org/10.1002/qj.4764>