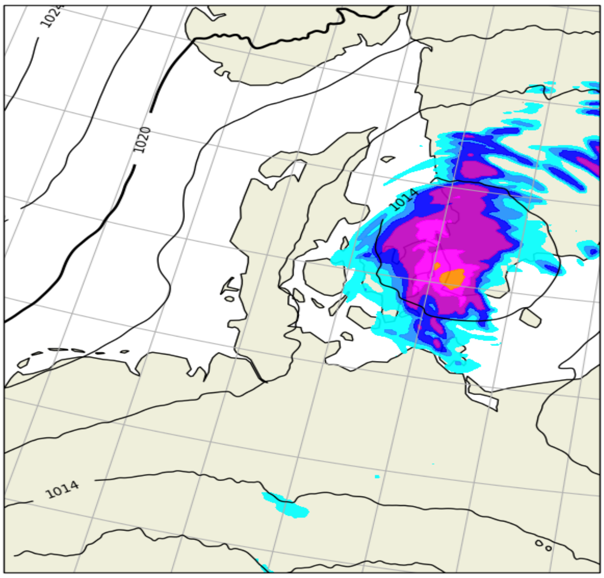
The Danish Meteorological Institute

DMI Report 21-31 Danish Regional Atmospheric Reanalysis

Final scientific report of the 2020 National Centre for Climate Research Work Package 3.2.1, Regional Reanalysis Pilot

DMI Report
15 January 2021

By: Xiaohua Yang, Kasper Stener Hintz, Carlos Peralta Aros, Bjarne Amstrup



The Danish Meteorological Institute

Colophon

Serial title	DMI Report
Title	DMI Report 21-31 Danish Atmospheric Reanalysis
Subtitle	Final scientific report of the 2020 National Centre for Climate Research Work Package 3.2.1, Regional Reanalysis Pilot
Author(s)	Xiaohua Yang, Kasper Stener Hintz, Carlos Peralta Aros, Bjarne Amstrup
Other contributors	
Responsible institution	Danish Meteorological Institute
Language	English
Keywords	Regional Reanalysis, HARMONIE
URL	https://www.dmi.dk/publikationer/
Digital ISBN	978-87-7478-705-1
ISSN	2445-9127
Version	15 January 2021
Website	www.dmi.dk
Copyright	DMI

The Danish Meteorological Institute

Content

1 Abstract 4
2 Resumé 5
3 Introduction 6
4 DANRA reanalysis system 9
5 DANRA production 12
6 DANRA performances in extreme weather events 18
7 Summary and outlook 22
8 Acknowledgement 24
9 References 25
10 Appendix DANRA output parameters 26
11 Previous reports 30

1 Abstract

A 2.5 km resolution, regional atmospheric reanalysis project, the DANish atmospheric ReAnalysis (DANRA), has been initiated at the Danish Meteorological Institute (DMI), within the framework of the Danish National Centre for Climate Research (Nationalt Center for Klimaforskning, NCKF). DANRA is aimed for production of a regional reanalysis for Danish area over an extended period of up to 70 years (1950 to 2020), providing Danish society with a high resolution, high fidelity atmospheric reanalysis dataset targeted for climate research and applications in Denmark and nearby regions.

The pilot phase of DANRA during 2020 has seen the establishment of the DANRA reanalysis system, which is based on the current operational HARMONIE, the Numerical Weather Prediction (NWP) system at DMI, with combination of model system infrastructure developed in the Copernicus Arctic Regional ReAnalysis (CARRA) project. Using the DANRA system, a reanalysis pilot has been produced covering 8 different time slices over the recent three decades, resulting in a total of 6-year worth of gridded dataset. In addition, a provisional reanalysis has been carried out for a selection of historical weather extremes in Denmark occurred in the last 40 years, facilitating a quality preview about the DANRA dataset. From an intercomparison against observations as well as the state of art Copernicus global reanalysis ERA5, it appears that the DANRA dataset is indeed of superior quality in reflecting local weather conditions including extremes, thus an ideal dataset for climate study. Much work has also been devoted to establish the system infrastructure on monitoring of DANRA production and on postprocessing of reanalysis data for different applications.

It is anticipated that in the coming years, DANRA production will be carried through backwards, with delivery of ca 12-18 years per year and aim to ultimately cover the full 70 year period as those by the global ERA5 reanalysis, the latter is used as lateral boundary in this reanalysis. More efforts will be needed in collection and utilisation of observation data in order to secure quality of reanalysis and to maximize the benefit of high-resolution reanalysis. Work will continue to explore the use of reanalysis time series for study about the recent climate extremes and their trends, using the gridded DANRA dataset. In addition, a near-real time DANRA stream will be established in order to facilitate Danish society with high quality datasets tracking the evolution of climate states in Denmark.

The Danish Meteorological Institute

2 Resumé

Danmarks Meteorologiske Institut (DMI) har påbegyndt et regionalt reanalyse projekt under Nationalt Center for Klimaforskning (NCKF), kaldet DANsk atmosfærisk ReAnalyse (DANRA). DANRA's målsætning er en produktion af en regional reanalyse for dansk område over en periode på 70 år (1950 til 2020) med en rumlig opløsning på 2.5 km. Dette giver det danske samfund et klimatisk datasæt i en hidtil uhørt høj opløsning, der kan benyttes til såvel forskning og applikationer i Danmark og omkringliggende regioner.

Pilotfasen af DANRA i løbet af 2020 har formået at opbygge den komplette modelopsætning for reanalysesystemet. Dette er baseret på HARMONIE, en moderne numerisk vejrmodel, der til daglig fungerer som operativ vejrmodel på DMI. Ved at benytte DANRA systemet, er der i løbet af 2020 blevet produceret 6 års reanalyse data fordelt over 8 forskellige perioder mellem 1993 og 2020.

Herudover er der foretaget foreløbige reanalyser for udvalgte historiske vejrekstremer i Danmark, der kan benyttes som foreløbig evaluering af kvaliteten af DANRA datasættet. Ved at sammenligne modeldata med observationer og andre reanalyse produkter, kan det foreløbigt konkluderes at DANRA datasættet er af en høj kvalitet. Det primære arbejde i 2020 har været dedikeret til at etablere systemets infrastruktur, monitoring af produktionskæden samt udvikling af post-processeringsværktøjer.

Det forventes at i de kommende år, vil DANRA produktionen fortsætte tilbage i tid og kunne levere ca. 12-18 års data per år. Målet er at dække den fulde 70-års periode som også dækkes af den globale reanalyse ERA5. ERA5 virker som laterale grænsebetingelser til DANRA. Yderligere ressourcer vil blive nødvendige for at indsamle og udnytte observationsdata for at få fuldt udbytte af denne højt opløselige reanalyse. Klimatendenser og klimaekstremer vil løbende bliver undersøgt i nærmere detaljer med det gridded DANRA datasæt. Herudover, vil et dataflow i nær real tid blive opbygget således at det danske samfund får adgang til reanalyse datasæt af høj kvalitet med mulighed for at overvåge nuværende klimatendenser i Danmark.

3 Introduction

The Danish National Centre for Climate Research (Nationalt Center for Klimaforskning, NCKF) has completed its first year in 2020. It has been a source of funding for the Danish Meteorological Institute and collaborators for climate change related research during this year. The 18 work packages fall under 4 general themes:

1. Arctic and Antarctic Research
2. Climate change in the near future
3. Use of climate data
4. Support for the IPCC

Work Package 3.2.1 is the pilot project for the DANish regional ReAnalysis (hereafter referred to as “DANRA”), which aims to produce a high resolution, high fidelity, gridded regional reanalysis dataset for Denmark and nearby waters over an extended period of up to 70 years. As the main objective during the start phase, the pilot project targets to build up the reanalysis system and the necessary infrastructure.

During recent decades, atmospheric reanalysis has emerged as a widely used tool for monitoring of climate states, their evolution in the recent past and the present day. In a reanalysis, past weather states such as wind, temperature, pressure, cloud and precipitation are reproduced on a regular grid using advanced data assimilation tools in Numerical Weather Prediction (NWP). However, different from NWP, the focus of reanalysis is reconstructing the past and not predicting the future. By using a fixed version, state of the art NWP model, and assimilating the most complete observation records and a perfect boundary data about historical period, a reanalysis dataset can deliver the most reliable, highest quality and consistent dataset about weather states in the recent past, providing an ideal time series to build statistics about the past and present climate states, and to use them as reference to monitor the future evolution. Statistics derived from reanalysis is also useful in developing forecast products about warning of extreme weather, such as those of Extreme Forecast Index (EFI), which enables the use of gridded historical record in warning about weather anomalies.

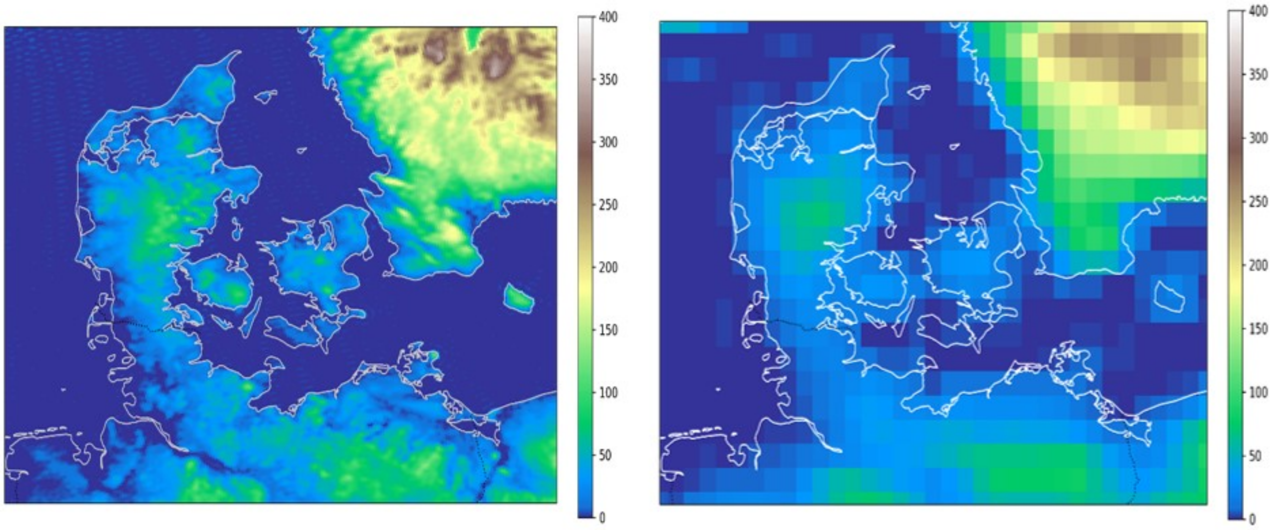


Figure 1. Orographic height in meters for regions around Denmark as represented by DANRA (2.5 km resolution, left) and ERA5 (31 km grid resolution, right)

Climate is an averaged weather state over a long period, and people experience climate and its changes from their daily encounter with weather. Most of the high impact weather occurs in local scales associated with underlying surface conditions such as orography, landscape and land use, which are far from homogeneous. Especially from the Danish perspective, despite the limited size and a relatively flat surface of the country, it is surrounded by water from all directions with long coast lines, creating conditions for weather features with considerable diversity and small scales. For climate study and adaptation, high resolution climate data sets reflecting local diversity are highly relevant and necessary. Over the last few decades, numerous global atmospheric reanalysis products have been developed, providing the global climate community with a gridded dataset for global climate. e.g, the EU-funded Copernicus global reanalysis ERA5 produced at ECMWF, is the state of art, highest resolution global reanalysis dataset (Hersbach et al, 2020). ERA5 is run on a 31 km grid resolution. However, reanalysis data with a resolution such as ERA5 is still too coarse on many applications when applied to study of climate issues in Denmark. With DANRA, we wish to fill in the gap by developing a higher resolution reanalysis dataset with focus on Denmark and nearby waters.

During 2020, the reanalysis system DANRA has been established using the operational Numerical Weather Prediction (NWP) model at DMI, HARMONIE, as system basis. After careful feasibility studies, an 800x600x65 grid mesh and 2.5 km resolution are selected as the model domain for the DANRA project, with a targeted ultimate production of 70-year reanalysis covering 1950 and 2020. Thanks to a very high resolution and a convection resolving model system (HARMONIE), DANRA reanalysis is anticipated to offer a clear added value to other reanalysis datasets with coarser resolutions. In Figure 1, orographic representation for Danish area by DANRA at 2.5 km grid is shown, together with those by ERA5 at 31 km grid, demonstrating a distinctively different representation of coastal landscape, which points to a clear potential advantage with DANRA in representation of near-surface climate states on temperature, wind, precipitation etc. Using DANRA, reanalysis was launched during the pilot phase in 2020, consisting of 8 time slices covering different periods between 1993 and 2019, producing a total of 6 year worth of reanalysis data for evaluation. In addition, provisional reanalysis for a selection of extreme weather events during the 1980s and the 2010s, mainly those major storms events and heavy summer convection events, have been conducted to examine the system fidelity to reproduce extreme weather conditions. Throughout DANRA production and evaluation of episode studies, ERA5 data have been used as a quality benchmark to verify added value of the high resolution DANRA reanalysis.

4 DANRA reanalysis system

4.1 DANRA system and model domain

Figure 2. DANRA reanalysis domain DKA, with a grid mesh of 800x600x65 and 2.5 km grid resolution. The boundary top is at ca 10 hPa.

Following the common strategy used in reanalysis projects such as ECMWF Re-Analysis (ERA) and Copernicus Arctic Regional Re-Analysis (CARRA), the present version of the NWP system used in routine weather forecast at DMI, DMI-Harmonie-40h1.1 (Yang et al 2017), has been selected as system baseline for DANRA. Harmonie 40h1.1, (Bengtsson et al 2017) is the state of the art non-hydrostatic, convection-permitting NWP system used in dozens of European national weather services participating in the ACCORD NWP research collaboration (A Collaboration on Convection scale Operational Model Research and Development) consortia (previously known as HIRLAM and ALADIN consortia). The main operation of the DANRA reanalysis consists of a 3-hourly analysis and forecast cycling, assimilating surface and upper air observation using HARMONIE 3-Dimensional VARiational data assimilation (3DVAR) scheme, followed by a short-range forecast of up to 18h. In view of logistic constraints, a model domain of DKA with 800x600x65 and 2.5 km grid distance has been selected (Figure 2). DKA used to be DMI's main operational model domain for Denmark and nearby waters during 2013 and 2017. For application in reanalysis, some of the system adaptations as implemented in CARRA (Yang et al 2020a), has been ported. These adaptations include use of hourly ERA5 reanalysis as the lateral boundary, use of the "brand" perturbation approach in derivation of background error statistics used for upper air data assimilation (Bojarova et al, 2019), as well as the system adaptations in reanalysis output. In data assimilation, conventional data as used by the ERA5 reanalysis (Hersbach et al, 2020) and the additional local observation data jointly collected by the Copernicus regional reanalysis projects CARRA and CERRA (Copernicus European Regional ReAnalysis), reprocessed atmospheric motion vector, satellite scatterometer wind and GPS radio-occultation bending angle data, have been used.

4.2 Boundary input

DANRA reanalysis applies hourly global ERA5 reanalysis as lateral boundary condition. Different from the situation in routine weather forecast in which global forecast is used as lateral boundary during forecast integration in a Limited Area Model (LAM) system, analysis boundary is used in regional reanalysis, providing an optimal lateral boundary that is otherwise unfeasible in a time-critical operational weather forecasting. As shown in Yang et al 2020b, use of reanalysis boundary helps a LAM system to maximize its benefit with high resolution, so that it becomes possible for a LAM reanalysis to achieve a superior skill in representation of atmospheric states for not only surface weather parameters, but also for quantities dominated by large scale, such as surface pressure and upper air key parameters. Similar to the CARRA approach as explained in Bojarova et al (2020), in preparation of structure function for DANRA reanalysis, a two step approach has been used to compute model error covariance statistics using two-time slices for selected summer (July 2012) and winter episodes (Jan 2017). Ensemble dataset of the ERA5 Ensemble of Data Assimilation (EDA) , with a 60 km resolution, has been used as lateral boundary condition in the ensemble setup for these computations. For surface sea states, gridded analysis data extracted from ERA5 has been used.

4.3 Enhanced observation data

DANRA assimilates near-surface observation of temperature, humidity at screen level and snow depth measurements. For upper air, in-situ (conventional) and satellite remote sensing data available from multiple sources are used. These include measurements of surface pressure collected via Global Telecommunication System (GTS) from SYNOP, SHIP and DRIBU network, wind over sea, profile measurement by radiosonde, aircraft and PILOT network. The baseline dataset used in DANRA reanalysis are those from ECMWF MARS archive for the input data used by the global reanalysis ERA5. In addition, where available, measurements data provided by the observation archive in the Nordic weather services of Denmark, Norway, Sweden and Finland are used, the latter through the joint collection by the two Copernicus regional reanalysis projects CARRA and CERRA, in which DMI has also been a contributing partner. It shall be pointed out that the observation data contributed by DMI on these datapools has been limited. Resource permitting, efforts are needed to convert the quality-assured observation data by DMI into the reanalysis for earlier years.

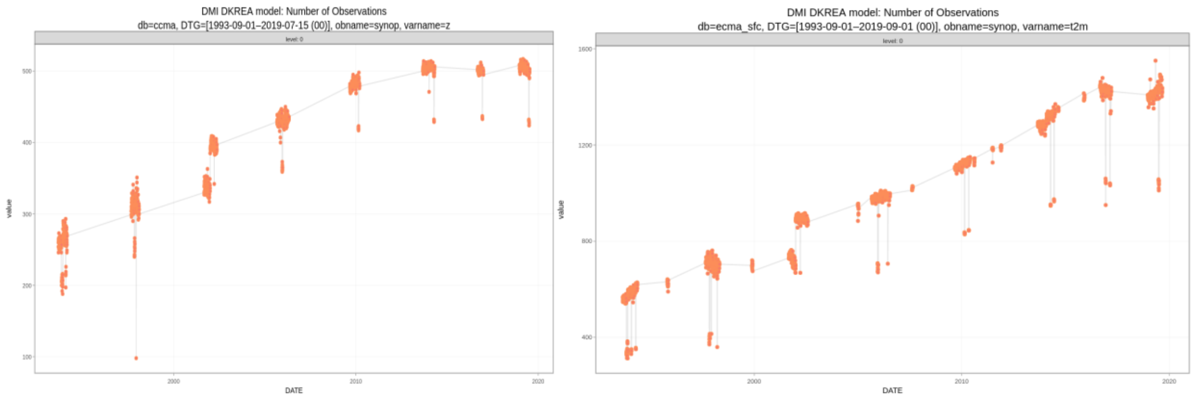


Figure 3. Assimilated observation report of surface pressure (left) and 2-meter temperature in DANRA during the assimilated DANRA periods between 1993 and 2020. It appears that the numbers of available in-situ near-surface observation data have increased over the period, at least for the DANRA reanalysis domain.

Substantial temporal variation exists in observation data coverage over the DANRA reanalysis period. Figure 3 plots data amounts of assimilated surface pressure and 2-metre temperature, illustrating a clear enhancement over years in terms of observation data amount used in DANRA. It is unclear to what extent this is attributed to an increased investment in observation networks in these regions, but part of the explanation could be due to an improved delivery of observation data by national weather services in recent decades, which enhanced the station density in the region. This is illustrated in Figure 4, in which 2-meter temperature observation distribution for a day in 1993 (Oct 1, 00 UTC) and 2019 (Jul 1, 2019) are shown. It shall also be pointed out that, implementation of automatic weather stations in the recent decades increased significantly temporal observation density so that for most of Europe, hourly surface observation became common.

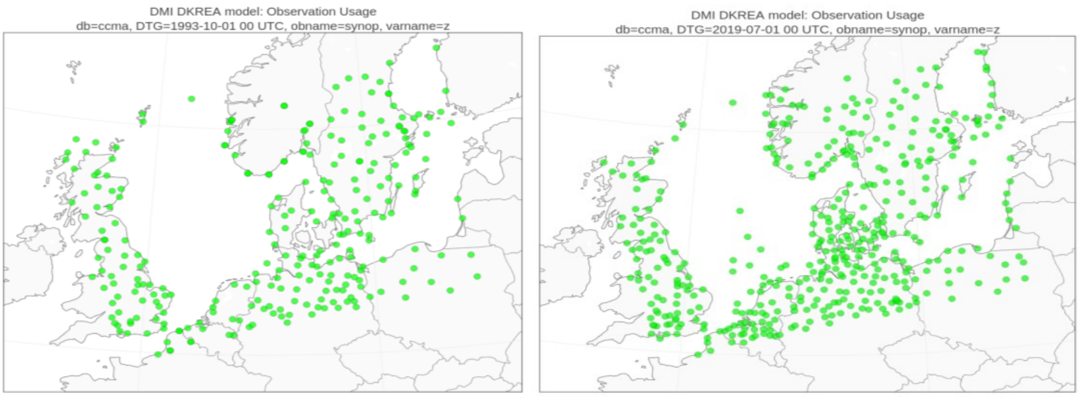


Figure 4. Assimilated surface pressure observation for 1 Oct 1993, 00 UTC (left) and 1 July 2019, 00 UTC. There appears to be an increase in the available surface observation for the DKA area during recent decades.

In recent decades, radiosonde observation data suffered a noticeable reduction in Europe. e.g, radiosonde measurements have been discontinued in Denmark. From Figure 5, the amount of radiosonde data shows a decline with time, whereas aircraft measurement starts to become a significant part of the observation network from the start of the century. As in CARRA, (Yang et al 2020a), a scheme to bias correct radiosonde temperature data has been implemented using ERA5 correction.

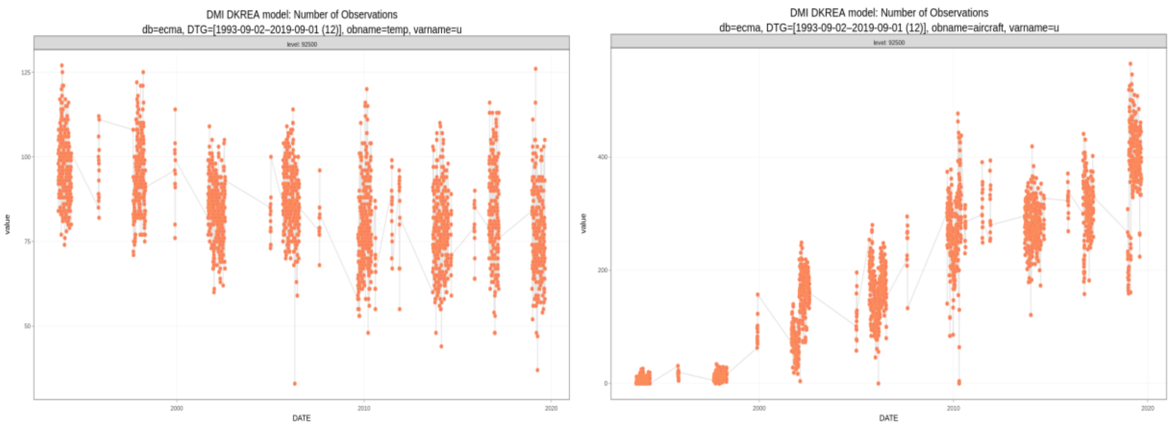


Figure 5. Assimilated numbers of upper air wind observation reports with radiosonde (left) and with aircraft measurement (right) in DANRA during the assimilation periods between 1993 and 2020. It appears that the numbers of available radiosonde data have experienced a decline over the period, whereas aircraft observation data increased during the same period.

Using an approach similar to CARRA reanalysis, DANRA assimilates a wide range of satellite observations, including those of atmospheric motion vector from EUMETSAT geostationary satellites, satellite microwave and infrared radiance, scatterometer and radio occultation data from EUMETSAT and NASA polar satellites, Variational bias correction is used for assimilation of radiance data. For blacklisting, CARRA approach has been adopted, utilising largely the relevant system infrastructure of the ERA5. DANRA assimilates bending angle from reprocessed GPS radio occultation data as provided by the Radio Occultation Meteorology Satellite Application Facility (ROM SAF), which is a decentralized operational RO processing centre under EUMETSAT. The RO data used in DANRA is the same used by CERRA and available at http://www.romsaf.org. As in CARRA, for scatterometer data, reprocessed climate data records (CDR) from EUMETSAT OSISAF for ERS Scatterometer, OceanSat/OSCAT, QuickScat/Seawinds and Metop/ASCAT-A for different periods are selected for an optimal and consistent data quality.

5 DANRA production

5.1 Organisation of production streams

Reanalysis is a resource-demanding exercise both in terms of High-Performance Computing (HPC) and staff resources. Selection of either ECMWF or DMI's own HPC platform has both pros and cons. In view of practical constraints and logistics, the pilot production is conducted on the ECMWF platform using the national HPC resource allocation for Denmark. Running DANRA at ECMWF HPCF makes it convenient to access ERA reanalysis boundaries and observation data archives. On the other hand, it is a substantial logistic issue to transfer selected reanalysis and monitoring data back to DMI platform for archiving and postprocessing, requiring careful data thinning.

Re-analysis stream	Reanalysis Period	Finished Period	Total number of finished months in 2020
1993	1994.09 - 1998.08	1993.09 - 1994.05	9
1997	1997.09 - 2002.08	1997.09 - 1998.04	8
2001	2001.09 - 2006.08	2001.09 - 2002.07	11
2005	2005.09 - 2010.08	2005.09 - 2006.07	11
2009	2009.09 - 2014.08	2009.09 - 2010.05	9
2013	2013.09 - 2017.08	2013.09 - 2014.06	10
2016	2016.09 - 2019.11	2016.09 - 2017.03	7

Re-analysis stream	Reanalysis Period	Finished Period	Total number of finished months in 2020
NRT	2018.12 - 2024.12	2018.12 - 2019.09	10

Table 1. DANRA reanalysis streams performed during the pilot phase in 2020, which consists of 8 streams covering different time slices for recent decades. DANRA is currently done at ECMWF HPCF, utilising allocated member state computation resources.

In terms of reanalysis configuration, DANRA runs intermittent HARMONIE data assimilation cycling with a 3-hour interval. For surface data assimilation, optimal interpolation scheme (OI) is used to assimilate screen level observation of temperature, humidity and snow depth, updating temperature and humidity information on surface and soil levels. For upper air, Harmonie 3DVAR is used to assimilate in-situ and remote sensing observation data as listed above, updating every three hour main prognostic variables such as wind, temperature, humidity and pressure. Note that not all the model states are updated via data assimilation. For hydrometeors, cloud properties, nonhydrostatic pressure departure, turbulent kinetic energy, etc, the background fields, i.e., 3-h forecast from previous assimilation cycle, are used. With output of data assimilation as initial condition, an integration of 3h in lead-time is made for all cycles, with exception of cycle 00 and 12 UTC, where a lead-time of 18h forecast is used instead. The longer forecast lead-time at 00 and 12 UTC is a design that enables the extraction of accumulated precipitation using model integration beyond a lead-time of 6h, in order to get rid of moisture spin-up often seen in the initial stage of model integration following assimilation. Such a configuration is similar to the normal practice used in other reanalysis projects such as ERA5 and CARRA.

Production of DANRA reanalysis is organised by 8 parallel streams, each of them covering a period of ca 5 years. Table 1 lists the streams launched in 2020 for time slices between 1993 and 2020. With the current strategy, the reanalysis is scheduled to be organised in several groups, each for about 15 to 20 years, with an anticipated production time of about one year per group. The production starts from the latest decades and proceeds backwards in time, with the target to finish the rest of the 70 years period in coming 4 to 6 years, depending on availability of resources. So far, 75 months of reanalysis, with about 6 years' worth of data, have been generated, serving as data pilot for evaluation. Examination of pilot data also enables necessary tuning before start of the final production.

It is estimated that a full year production of DANRA consumes ca 8,000 K System Billion Unit (SBU), thus a 12-year reanalysis would consume ca 50% of the annual SBU allocation to Denmark on the ECMWF HighPerformanceComputing Facility (HPCF).

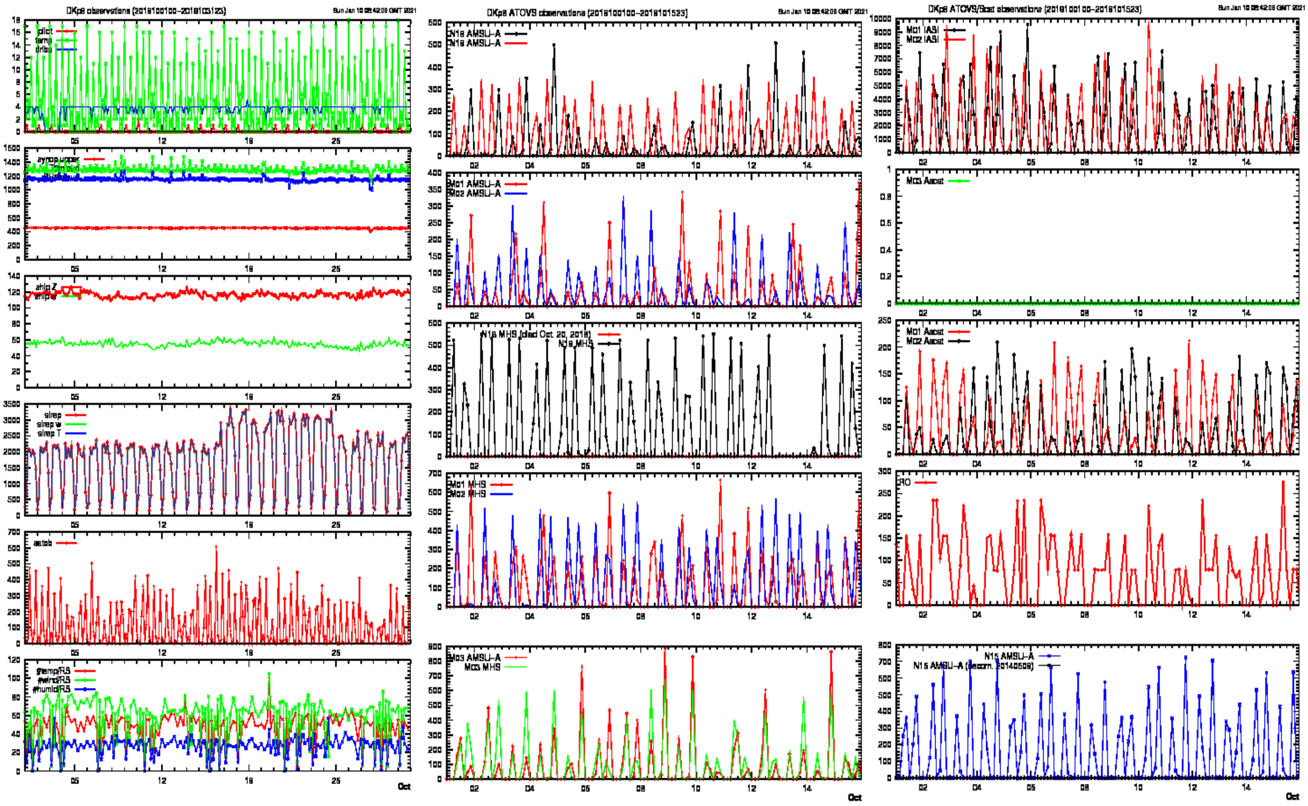


Figure 6. Time series of daily sum of assimilated observation report for the month of Oct 2019. Left, conventional observation. Middle and Right: satellite remote sensing data with radiance, scatterometer and Radio Occultation bending angle data.

5.2 Reanalysis data

Reanalyses generate a huge amount of gridded model data, which is a logistic challenge for storage. In this work, an extraction procedure has been developed, in which hourly data for some of the commonly used reanalyses parameters are extracted from each reanalysis cycle, and archived in GRIB format. The selection of parameters has drawn experiences from reanalysis projects of CARRA and MERA (MetEireann ReAnalysis, Eoin Whelan, Met Eireann, personal correspondence).

DANRA reanalysis is defined to consists of following hourly data,

a) Parameters valid at analysis time with a 3-h interval
b) Parameters valid at other time points outside of the 8 analysis time points ("asynoptic time"), which are provided by 1-h and 2-h forecasts
c) As an exception, hourly precipitation fields are derived from the difference of accumulated forecast between two consecutive forecasts lead time that is one hour apart. In order to minimise effect of moisture spin-up, model data used to derive precipitation are those with a minimum lead-time of 6-h.

Some of the parameters are defined on height or pressure levels, rest are on single level.

A detailed list of selected output parameters can be found in the Appendix, which also provides information about units and other meta-data information.

5.3 Monitoring and evaluations

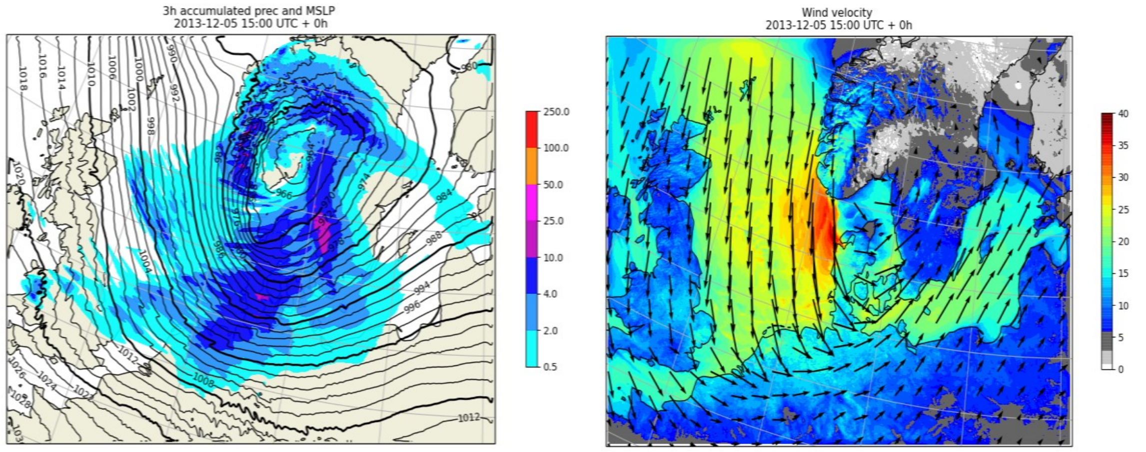


Figure 7. Storm Bodil as captured in DANRA, with Mean Sea Level Pressure and 3 h accumulated precipitation in units of kg/m² (left) and wind speed (right) at 10 m in units of m/s on 5 Dec 2013, 15 UTC.

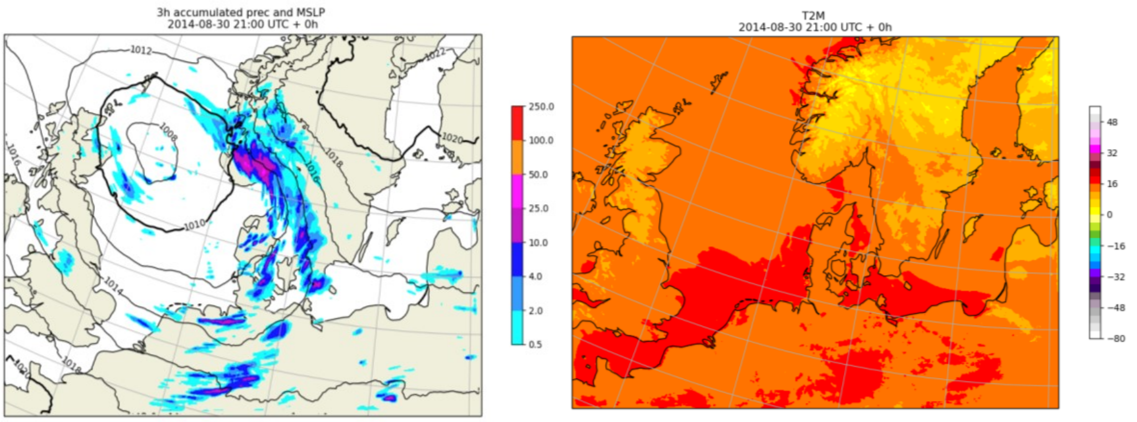


Figure 8. Cloudburst event in Copenhagen on Aug 30 2014, 21 UTC as simulated by DANRA, with 3h accumulated precipitation and Mean Sea Level Pressure (left) and temperature at 2 m (right).

For daily monitoring of DANRA production, tools developed for HIRLAM-C research programme and for CARRA reanalysis (Yang et al, 2020) have been adapted, including the utility package GL and MONITOR, as well as a web interface. With these, daily production throughput and Gantt chart, statistics on observation amounts used for assimilation, monitoring of data assimilation properties (background-observation, analysis - background, bias statistics etc.), observation verification and near-surface and upper-air parameters including intercomparison against ERA5, and weather charts for key parameters, are presented on a web interface for easy, regular monitoring and detection about various types of anomaly in production.
Figure 6, e.g., shows an example time series of assimilated total observations of different types for the month of Oct 2019.

Figure 7 and 8 show examples of weather charts automatically generated every 3 hours for key weather parameters. In Fig 7, DANRA simulation for Mean Sea Level Pressure, precipitation and wind during the category 4 hurricane “Bodil” on Dec 5 2013 is shown. In Fig 8, the weather situation for the major Copenhagen cloudburst event on Aug 30 2014 is shown. During that event, convective precipitation of over 100 mm was observed at several locations in central Copenhagen.

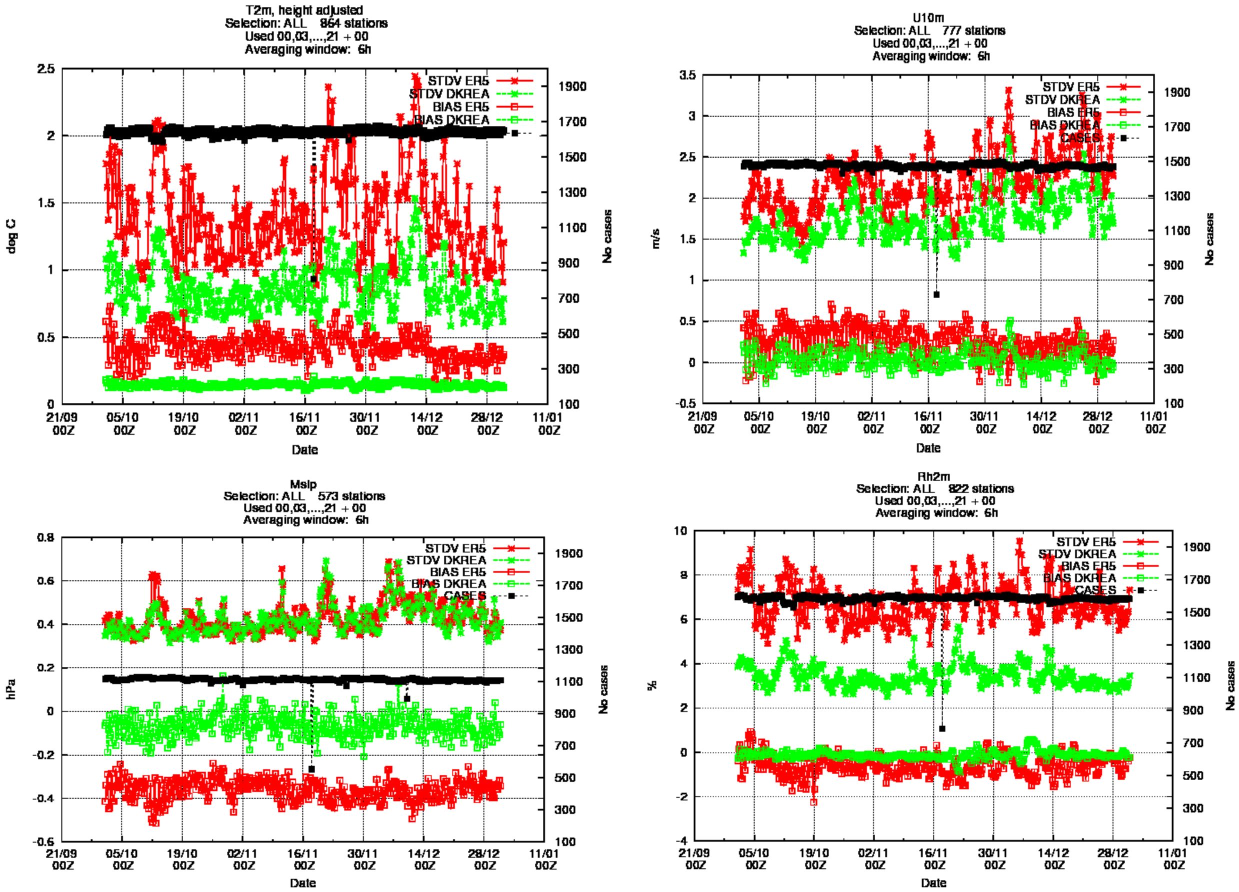


Figure 9. Time series of the averaged standard deviation errors (STD) and bias errors (Y-axis) with DANRA reanalysis (in green) validated by surface observations, in comparison to those by ERA 5 (in red) for 2 m temperature (upper left), 10 m wind speed (upper right panels), Mean sea level pressure (lower left) and relative humidity at 2 m (lower right), for the period between Oct 1 and Dec 31 2013. Dark crosses in the figures are the total number of observation data available for intercomparison.

In monitoring of DANRA reanalysis production, verification information extracted from ERA5 is regularly intercompared, in which ERA5 is used as a quality benchmark, partly to confirm the added value of the high resolution regional reanalyses, partly to help detection of abnormal situations in the DANRA reanalysis potentially associated with technical issues.

Through regular monitoring of observation validation and intercomparison to ERA5, added values of the regional reanalysis with DANRA can be closely followed for different production streams and periods. Figure 9 and Figure 10 illustrate, e.g., the verification of screen level parameters comparing reanalyses of DANRA and ERA5 against synoptic observation data for two 3-month period, one for the winter months between Oct and December 2013 (with two hurricane events included), another for the summer months of June-August 2014 (with numerous convective precipitation events). With a generally much smaller error level in temperature, wind, surface pressure and relative humidity, DANRA is shown to have a clearly superior fidelity in representing climate status. Among the key parameters as shown in the figures, especially the DANRA fit to observations in temperature and humidity are substantially better than ERA5 thanks to surface data assimilation and a much higher grid resolution. For wind speed over land, which is not assimilated, DANRA, DANRA is in general also more superior than ERA5 in the fit to observation, especially in standard deviation error, thanks to the resolution advantage. For Mean Sea Level pressure, the quality is less dominated by model resolution, hence the advantage with DANRA is not equally substantial.

DANRA production is currently still in the pilot phase, in which production streams are being spin-up from a cold start. Between consecutive production streams, an one year overlap is arranged to account for the spin-up time needed to ‘warm up’ soil parameters from a cold start. Such an overlapping period also serves purposes to cold-start variational bias correction coefficients, which normally takes a few months to reach a steady-state for optimal assimilation of satellite radiance data. As such, the reanalysis data collected during the pilot phase, which mostly covers the initial cold-start period for each production stream, are of the nature of provisional data. Based on experiences in CARRA, data in the spin-up period generally match well with those of final dataset with few exceptions directly related to soil parameters and as such, are relevant in the general evaluation of reanalyses.

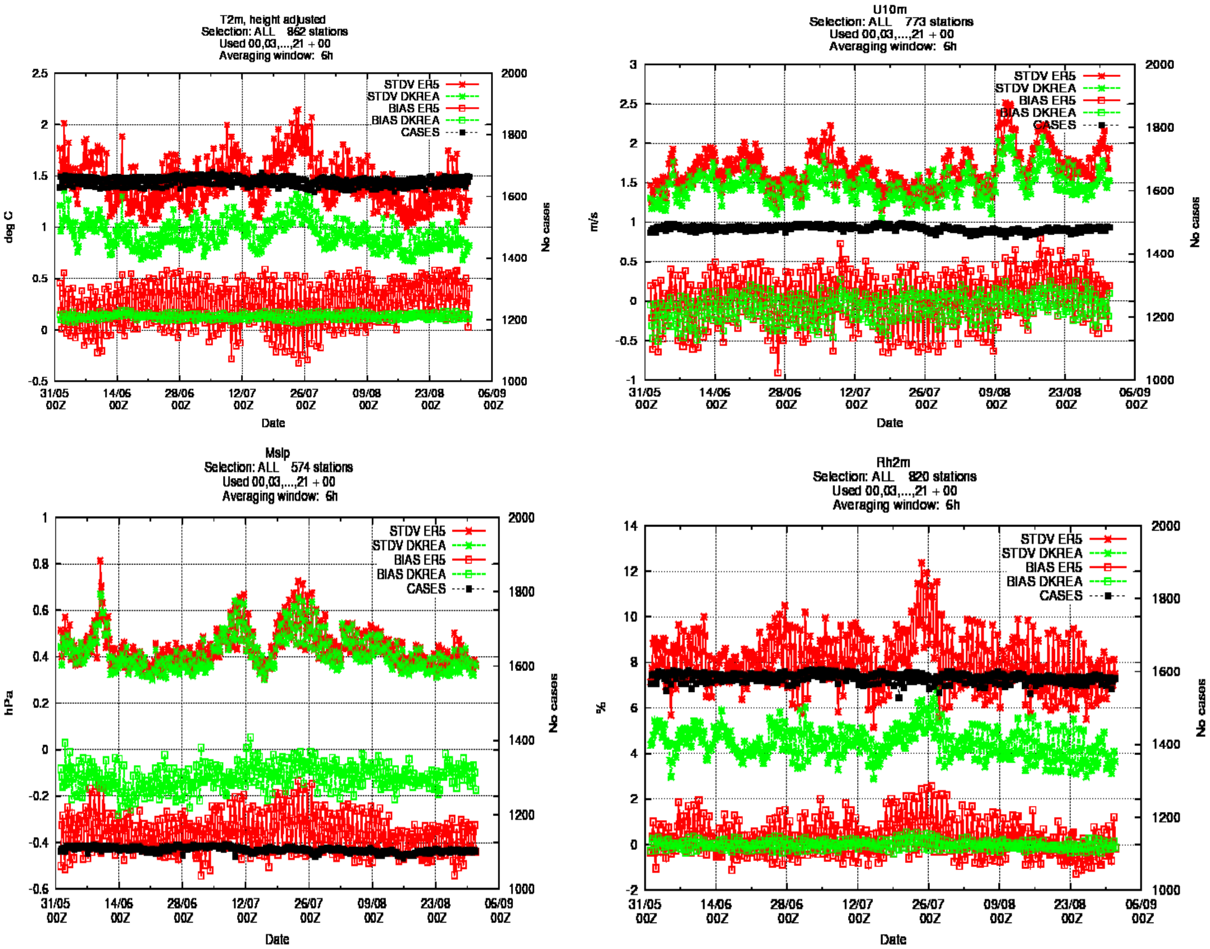


Figure 10. same as Fig 9 but for the summer period between June 1 and Aug 31 2013.

6 DANRA performances in extreme weather events

Recent decades have seen increasingly more weather extremes around the world, and the trend seems to accelerate due to global warming. For reanalysis products, it is of critical importance to examine its fidelity in reproducing major weather events during the reanalysis period.

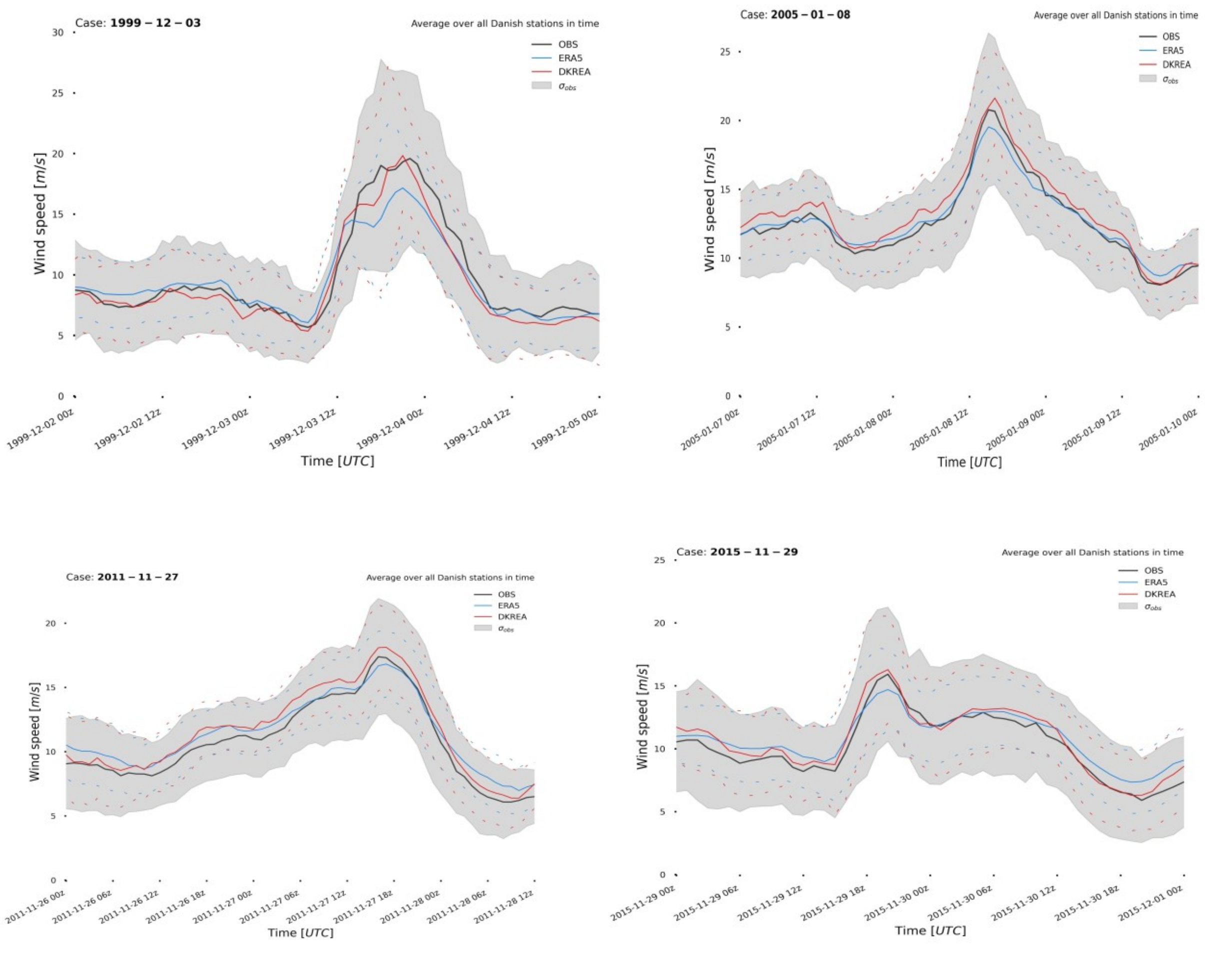


Figure 11. Time series of wind speed averaged over standard Danish surface stations during storm episodes, with black color depicting observations, red color for DANRA, blue for ERA5. Variances in each data series are represented by grey shading. Upper left for Dec 3 1999, upper right for Jan 8 2005, lower left for Nov 27 2011, lower right for Nov 29 2015 (storm ‘Gorm’). Both of the DANRA and ERA5 reanalysis appear to reproduce well the evolution of the storm events, with DANRA clearly superior in reproducing the peak value in wind speed.

6.1 Major historical Danish storms

During the DANRA pilot phase, reanalyses for the most damaging Danish storms recorded in the recent decades have been examined as a quality preview about the DANRA data. In Figure 11, time series of wind speed averaged over standard Danish surface stations during the 4 major storm episodes in 1999, 2005, 2011 and 2015 are shown. The observation data are those from the standard SYNOP stations in Denmark. Both DANRA and ERA5 reanalyses seem to reflect well the evolution of these storms, but the fit with DANRA to peak values are especially impressive.

		Lowest central pressure [hpa]			Highest mean wind [m/s]		
	Category of Storm Intensity	**OBS**	**DANRA**	**ERA5**	**OBS**	**DANRA**	**ERA5**
1981-11-24	**4**	968.0	968.3	967.9	35.0	28.5	26.2
1983-01-18	**4**	968.8	964.1	965.1	31.4	27.4	23.2
1990-02-26	**4**	951.9	949.5	949.7	31.0	28.6	23.6
1991-01-09	**4**	973.4	969.0	968.4	33.0	29.9	24.5
1993-01-14	**3**	976.2	975.6	975.0	30.0	31.3	27.2
1999-12-03	**4**	952.4	955.0	958.5	41.2	39.5	27.5
2005-01-08	**3**	965.1	965.0	966.8	32.0	32.0	28.3
2011-11-27	**2**	978.8	979.3	973.2	28.0	27.6	23.4
2013-10-28 "Alan"	**4**	967.5	965.9	966.9	38.0	33.7	23.6
2013-12-05 "Bodil"	**4**	966.0	966.7	965.5	31.4	31.2	25.8
2015-11-29 "Gorm"	**3**	971.7	971.9	971.4	34.0	32.9	23.3

Table 2. Central low pressure and peak mean winds as observed and simulated in the DANRA and ERA 5 reanalyses. The numbers in blue color are assigned to the dataset that fits best to those observed. The listed classification of storm intensity is based on the storm list issued by DMI (Cappelen, 2011). Note that starting from Oct 2013, storm names are assigned by DMI, which are specified in the first column of the table.

Table 2 lists the validation on DANRA data on central low pressure and peak wind speed, for each of the most severe Danish storms in the last 4 decades, many of those with observed peak winds surpassing hurricane scale and labeled as category 3 or 4 storms. Also listed for comparison are the corresponding extractions from ERA5. For readers convenience, the model values that are closest to observed ones are marked with blue color. As shown in the table, DANRA reproduces, in general, rather well these strong storm events, with simulated peak wind quite much closer to the observed extremes than the counterparts in ERA5. In comparison, the wind maxima from ERA5 generally tend to be too weak to reach hurricane scale, presumably due to limitation in horizontal resolution. On the other hand, between DANRA and ERA5, the relative skill in simulated central low pressure for these storm situations are rather mixed, indicating the dominance of large scale on such property, for which the role of model resolution may not be decisive.

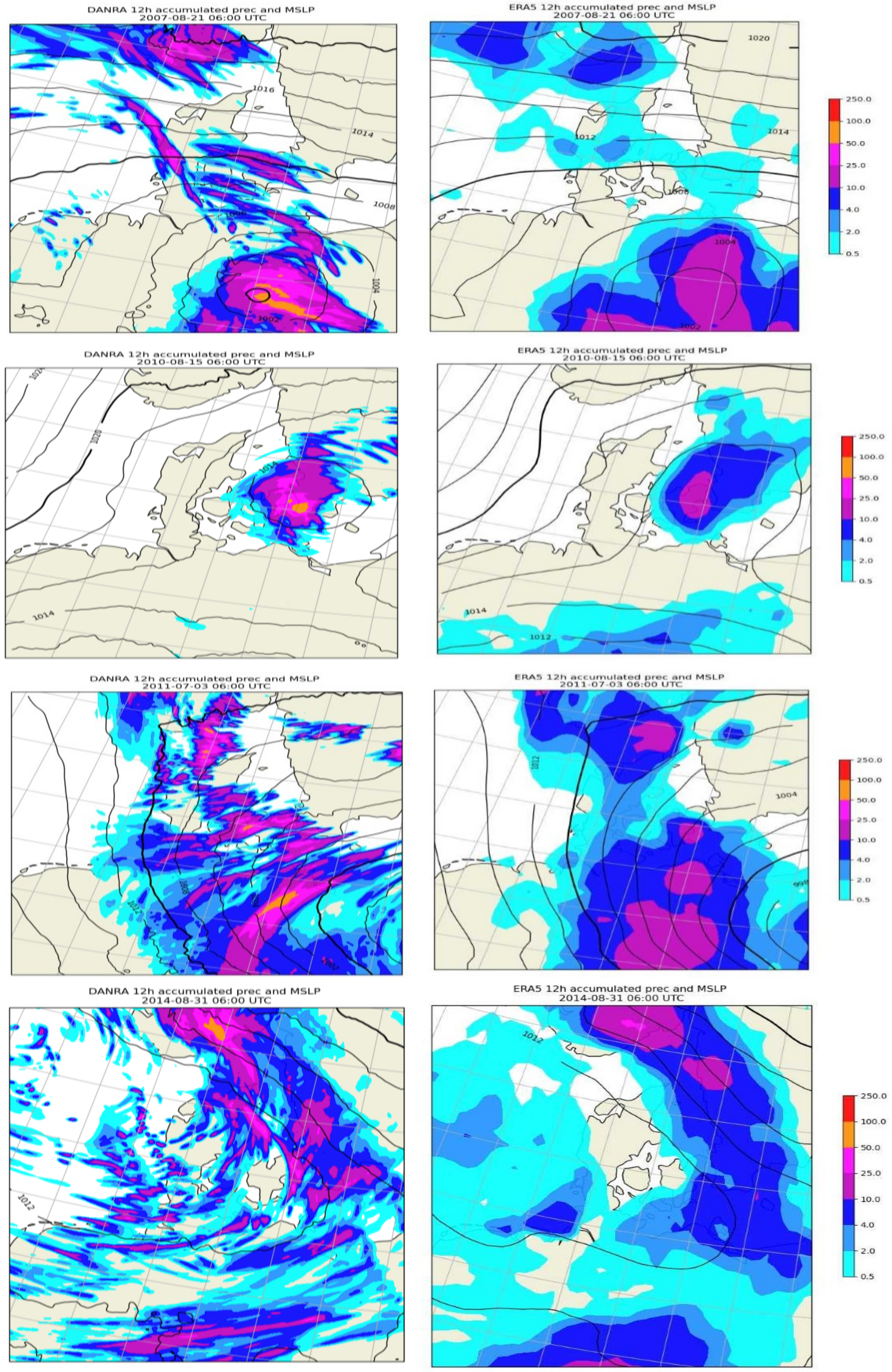


Figure 12. Representation by reanalysis DANRA (left) and ERA5 (right) for the heavy flashflood events in Denmark. All of the selected extreme convection cases occurred during the last two

decades, each with observed precipitation totalling over 100 mm. The plotted are the 12h accumulated precipitation up to the time points as labeled in the figure captions. From top to bottom, the cloudbursts in south Jutland on Aug 20, 2007, Copenhagen city on Aug 14 2010, July 2 2011 and Aug 30, 2014.

It shall be pointed out that, presently, observation station lists used in these provisional DANRA case studies are limited to those standard WMO stations, leaving some of the regular observation stations out. This has the consequences that, for some storm cases, the maximum values listed in the above may not be consistent with the official record at DMI. In fact, this is the case with storm Bodil on Dec 5 2013, in which the highest recorded mean wind was 36.6 m/s as measured at station Thorsminde, which is quite a bit higher than the above-listed maxima. The observation data from the station is normally available to DMI but it is outside of GTS exchange due to a different station ownership.

6.2 Major cloudburst events

Past decades have seen increasingly many occurrences of extreme convective precipitation, often with limited horizontal scales and short duration. Further, phenomena like extreme convective precipitation are also projected to become more frequent against a warming climate, raising concerns on more severe consequences to human life and properties. Forecasting such events is a major challenge in weather forecasting. Most importantly, representation of convective precipitations is quite demanding for NWP models especially in terms of resolution, model dynamics and physics, assimilation of relevant observation data. With the DANRA system, thanks to a very fine resolution of 2.5 km and an advanced cloud-resolving model HARMONIE; there shall be a better opportunity for the reanalysis system to describe convective precipitation events in contrast to other coarser resolution reanalyses. We examine here the performances of DANRA for four of the most extreme summer convection cases selected from the collected cases for extreme precipitation events in recent decades (Cappelen, 2020).

The first of these occurred on Aug 20, 2007 near Danish-German border in Southern Jutland, with accumulated total precipitation of over 100 mm near Gråsten, according to unofficial observation and radar derivals (Flemming Vejen, personal communication). The heavy precipitation amounts are associated with a series of consecutive mesoscale cells. The downpouring cloudburst in a very short time even washed away a segment of railway track. From Figure 12, a narrow stripe of convective precipitation band is visible across southwest Jutland from the DANRA simulation, corresponding well to the observation. In contrast, little signal for heavy convection is visible in the corresponding ERA5 figure. The remaining three selected events as shown in Fig 12 all occurred in the Copenhagen Metropolitan area (on Aug 14 2010, July 2 2011 and Aug 30 2014, respectively), with peak precipitation of over 100 mm, resulting in severe flooding in the city. For all the three events, DANRA appear to have simulated rather well these strongly convective events, with indication of small scale extremes that clearly surpass the criteria for cloudburst warning in Denmark, (>15 mm in half hour), although with an insufficient sharpness and magnitude in comparison to the observed extreme. For these events, ERA5 has a good signal for the Aug 14 2010 convection albeit weaker than that by DANRA. For the events on July 2 2011 and Aug 30 2014, ERA5 lacks representation.

7 Summary and outlook

Atmospheric reanalysis has, over the past decades, emerged as an important tool in climate study and monitoring. In a reanalysis, the data assimilation tool in the operational NWP system is used to produce long range, gridded time series for the past weather states. Compared to climate statistics based on observation data, a reanalysis dataset has an obvious advantage with a gap-free, continuous and gridded dataset, and is especially superior in temporal consistency due to the use of a fixed model version. In this work, we establish the 2.5 km atmospheric reanalysis system DANRA, with a target to produce reanalysis time series for Danish area over an extended range of up to 70 years, providing Danish society with an unprecedented high resolution, high fidelity reanalysis product to aid climate study in Denmark.

During the initial phase, the DANRA reanalysis system and its production infrastructure have been established, including those for monitoring, quality assurance and postprocessing. The reanalysis system combines features in the operational NWP (DMI-HARMONIE 40h1.1) and the regional reanalysis system (CARRA). For logistic reasons, DANRA pilot is conducted at ECMWF HPCF platform, making it convenient to utilize some of the system infrastructure for the Copernicus regional reanalysis projects CARRA and CERRA such as observation data. Provisional DANRA reanalyses, consisting of 8 simultaneous production streams, have been conducted, generating 6 year worth of reanalysis pilot data. For quality assurance, a monitoring infrastructure is being established to follow production of the reanalysis, including development of utilities to generate and monitor relevant time series. In this connection, time series of GTS observation and ERA5 for all the targeted reanalysis period have been extracted to enable a timely intercomparison and detection of anomaly.

Basic monitoring and quality assurance of the DANRA data is done by a regular validation to in-situ observation data. The corresponding ERA5 data are utilised here as a quality benchmark. This is based on the assumption that, for key weather parameters such as near surface temperature, wind and humidity, high resolution regional reanalysis shall normally over-perform those by the global reanalysis due to a substantially higher resolution, thus a better fundamental to resolve many local effects due to the underlying surface conditions. The validation intercomparison for the DANRA and ERA5 dataset during the pilot period, as presented in this report, indeed confirms this, reaffirming the added values of the high resolution reanalysis to represent climate states. Regular intercomparison of DANRA data to ERA5 is also a convenient and efficient way in helping detection of an eventual DANRA production anomaly. e.g., if the verification time series for DANRA suddenly deviates substantially from a normal error level and if such is not observed also for ERA5, it may suggest some errors in the DANRA production suites, such as a loss of critical observation data stream in data assimilation, etc.

For climate time series airmed for use in assessment of climate status and trends, it is of critical importance that such is able to represent weather extremes. With global warming, weather extremes are anticipated to occur more frequently and with increased magnitude. For Denmark, winter storms and summer cloud bursts events are two of the most common high impact weather phenomena, both of these are strongly affected by local surface conditions, hence requiring high model resolution. Especially for summer convective weather such as cloud burst and heavy precipitation, observed extremes often are with rather short spatial and time scales. In development of the DANRA pilot, special attention has been devoted to examine its performance in representing extreme storms in autumn and winter, and the cloud burst events in summer. Dozens of the most severe storm cases in the last 40 years, most of them at hurricane scales, have been examined. It is concluded that DANRA reanalysis, thanks to its high resolution, is clearly more accurate in reproduction of the strong wind conditions in comparison to ERA5 dataset. For the most severe convective precipitation extremes as observed in the last two decades, DANRA is also found to be able to reproduce cloudbursts in all of the 4 examined events, whereas the ERA5 reanalysis either are too weak or fail to reproduce. Clearly, the cloud-resolving HARMONIE model system has been essential to represent the challenging severe convection situation, making it a superior tool for use in regional climate study for Danish area.

In recent years, it is increasingly common for the climate research community to use Near-Real Time (NRT) reanalysis, to generate regular updates about latest climate status and trend, with reference to the multi-year statistics built on a long time series. One such example is the interim reanalysis time series ERA-interim, which has recently been upgraded to ERA5-NRT. ERA5-NRT is typically with a 5 day delay (Hand Hersbach, personal communication) . For DANRA, one of the production streams, with a cold start on Dec 20 2018, is anticipated to reach near-real time during the first half of 2021. Afterwards, the suite can be continuously maintained, with a production delay of ca 10 days. It is anticipated that, in about one year, when the DANRA reanalysis is extended to be sufficiently long (e.g., > 30 years), it becomes feasible for DMI to issue regular updates about the latest climate status for Danish area. Statistics derived from reanalysis will also be useful to develop forecast products about warning of extreme weather, such as those of Extreme Forecast Index (EFI), which enables the use of gridded historical record in warning about weather anomalies.

8 Acknowledgement

The DANRA pilot is supported by the Danish “National Center for Klimaforskning (NCKF)” and one of the work packages in the NCKF during 2020. Kai Sattler contributed to preparation work at start phase in porting of reanalysis system to local computer platform at DMI. DANRA has also benefited from the system and observation data infrastructure in the EU-funded Copernicus global reanalysis (ERA5), Arctic Regional Reanalysis (CARRA) and Copernicus European Regional Reanalysis (CERRA). Eoin Whelan is acknowledged for providing GRIB data extraction script from Met Eireann ReAnalysis (MERA). EUMETSAT ROM-SAF team is acknowledged for making available re-processed Radio Occultation bending angle data for this study.

9 References

Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., de Rooy, W., Gleeson, E., Sass, B. H., Homleid, M., Hortal, M., Ivarsson, K.-I., Lenderink, G., Niemelä, S., Nielsen, K. P., Onvlee, J., Rontu, L., Samuelsson, P., Santos Muñoz, D., Subias, A., Tijm, S., Toll, V., Yang, X., Køltzow, M. Ø., 2017: **The HARMONIE-AROME model configuration in the ALADIN-HIRLAM NWP system**, Monthly Weather Review, May 2017.

Bojarova J, Schyberg H., Dahlgren P., Køltzog M., Yang, X., Peralta C , 2020: **Uncertainty estimation method.** Copernicus Climate Change Service contract 2018/C3S_D322_Lot2_METNorway/SC2 Deliverable Report no C3S_D322_Llot2.1.1.2.

Cappelen John, 2011, **Storm og ekstrem vind i Danmark**, DMI Technical Report 11-12. (https://www.dmi.dk/fileadmin/Rapporter/TR/tr11-12.pdf)

Cappelen John, 2020, **Ekstreme nedbørhændelser i Danmark - opgørelser og analyser til og med 2019**, DMI Technical Report 20-06. (https://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2020/DMIRap20-06.pdf)

Hersbach H. and co-authors, 2020, **The ERA5 global reanalysis**, Quarterly J. of Meteorol. Soc. Vol 146, Issue 730, 1999-2049.

Yang, X., B.S.Andersen, M. Dahlbom, B. H. Sass, S. Zhuang, B. Amstrup, C. Petersen, K. P. Nielsen, N. W. Nielsen, A. Mahura, 2017: **NEA, Operational implementation of HARMONIE 40h1.1 at DMI,** ALADIN-HIRLAM Newsletter 8, 2017.

Yang, X., Nielsen K.P., Dahlbom, M., Amstrup, B., Peralta, C., Høyer, J., Englyst, P.N., Schyberg, H., Homleid, M., Køltzow, M.A.Ø., Randriamampianina, R., Dahlgren, P, Støylen, E., Valkonen, T., Palmason, B., Thorsteinsson, S., Bojarova, J., Körnich, H., Lindskog, M., Box, J., Mankoff, K., 2020a: **C3S Arctic regional reanalysis – Full system documentation**. Copernicus Climate Change Service contract 2018/C3S_D322_Lot2_METNorway/SC2 Deliverable Report no C3S_D322_Lot2.1.1.1.

Yang, X., Nielsen K.P., Dahlbom, M., Peralta, C., Schyberg, H., Homleid, M., Køltzow, M.A.Ø., Randriamampianina, R., Dahlgren, P., Vignes, O., Støylen, E., Valkonen, T., Bojarova, J., Lindskog, M., Hagelin, S., Körnich, H., Palmason, B., Thorsteinsson, S., 2020b: **CARRA report on test and verification**. Copernicus Climate Change Service contract 2018/C3S_D322_Lot2_METNorway/SC2 Deliverable Report no C3S_D322_Llot2.1.2.

10 Appendix DANRA output parameters

Name	Units	Level type	Level
AROME hail diagnostic	kg m**-2	surface level	0
CAPE out of the model	J kg-1	surface level	0
Cloud base	m	surface level	0
Cloud ice	kg m**-2	pressure level	100,200,250,300,400,500, 600,700,800,850,900,925,10
Cloud top	m	surface level	0
Cloud water	kg m**-2	pressure level	100,200,250,300,400,500, 600,700,800,850,900,925,10
Direct normal irradiance	W m**-2	surface level	0
Geometrical height	m	20	27315
Geometrical height	m	5	0
Geopotential	m**2 s**-2	pressure level	100,200,250,300,400,500, 600,700,800,850,900,925,10
Global radiation flux	W m**-2	surface level	0
Graupel	kg m**-2	surface level	0
Gust u-component	m s*-1	surface level	10
Gust v-component	m s*-1	surface level	10
High cloud cover	(0 - 1)	surface level	0
Icing index	-	surface level	0

Name	Units	Level type	Level
Latent Heat Sublimation	J kg**-1	surface level	0
Latent heat flux through evaporation	W m**-2	surface level	0
Long-wave radiation flux	W m**-2	surface level	0
Low cloud cover	(0 - 1)	surface level	0
Maximum temperature	K	surface level	2
Medium cloud cover	(0 - 1)	surface level	0
Minimum temperature	K	surface level	2
Mixed layer depth	m	surface level	0
Momentum flux u-component	N m**-2	surface level	0
Momentum flux v-component	N m**-2	surface level	0
Net long-wave radiation flux (atmosph.top)	W m**-2	surface level	0
Net long-wave radiation flux (surface)	W m**-2	surface level	0
Net short-wave radiation flux (atmosph.top)	W m**-2	surface level	0
Net short-wave radiation flux (surface)	W m**-2	surface level	0
Precipitable water	kg m**-2	surface level	0
Precipitation Type	-	surface level	0
Pressure	Pa	103	0
Pressure	Pa	surface level	0

Name	Units	Level type	Level
Pseudo satellite image: cloud top temperature (infrared)	-	surface level	0
Pseudo satellite image: cloud water reflectivity (visible)	-	surface level	0
Pseudo satellite image: water vapour Tb + correction for clouds	-	surface level	0
Pseudo satellite image: water vapour Tb	-	surface level	0
Rain	kg m**-2	surface level	0
Relative humidity	%	pressure level	100,200,250,300,400,500, 600,700,800,850,900,925,10
Relative humidity	%	surface level	2,30,50,75,100,150,200, 250,300,500
Sensible heat flux	W m**-2	surface level	0
Short-wave radiation flux	W m**-2	surface level	0
Snow Sublimation	kg m**-2	surface level	0
Snow	kg m**-2	surface level	0
Temperature	K	pressure level	100,200,250,300,400,500, 600,700,800,850,900,925,10
Temperature	K	surface level	0,2,30,50,75,100,150,200, 250,300,500
Total cloud cover	(0 - 1)	surface level	0
Fog cover	(0 - 1)	surface level	2
Total precipitation	kg m**-2	surface level	0

Name	Units	Level type	Level
Vertical velocity	m s**-1	pressure level	100,200,250,300,400,500,60 700,800,850,900,925,950,10
Visibility	m	surface level	0
Water equivalent of accumulated snow depth	kg m**-2	surface level	0
Water evaporation	kg m**-2	surface level	0
u-component of wind	m s**-1	pressure level	100,200,250,300,400,500,60 700,800,850,900,925,950,10
u-component of wind	m s**-1	surface level	10,30,50,75,100,150,200, 250,300,500
v-component of wind	m s**-1	pressure level	100,200,250,300,400,500,60 700,800,850,900,925,950,10
v-component of wind	m s**-1	surface level	10,30,50,75,100,150,200, 250,300,500

11 Previous reports

Previous reports from the Danish Meteorological Institute can be found on:
https://www.dmi.dk/publikationer/