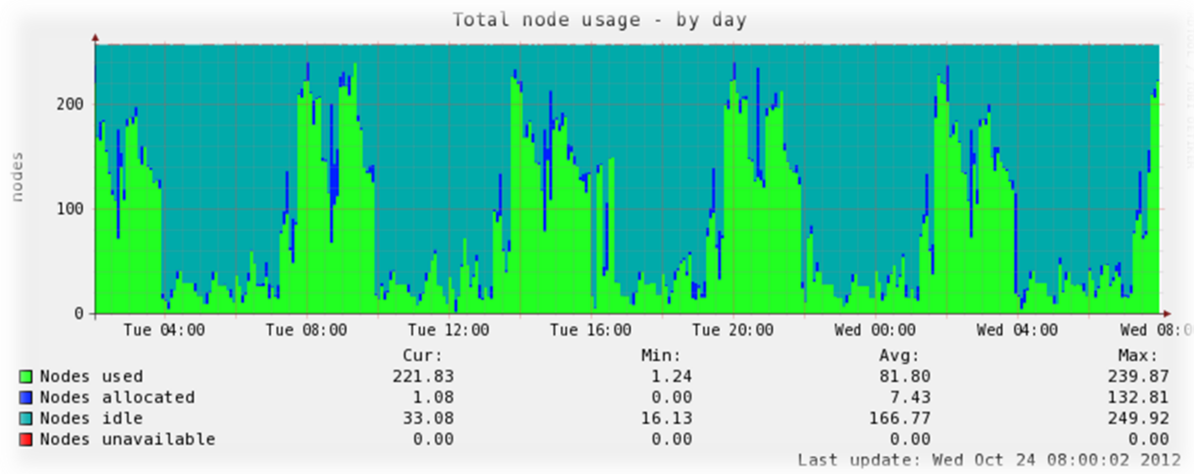
Dmi
Danish Ministry of Climate, Energy and Building

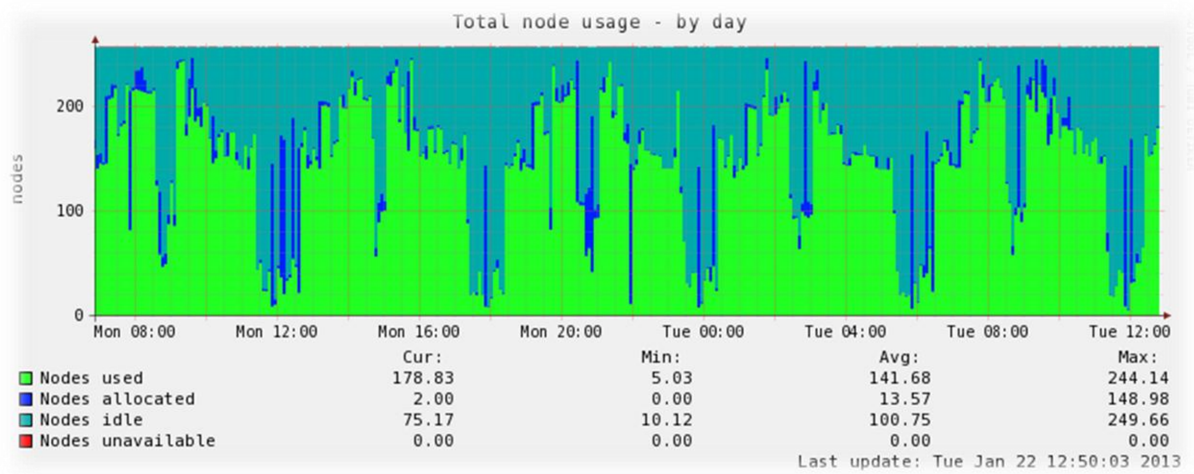
Technical Report 12-21

DKA37: HARMONIE Upgrade in December 2012

Xiaohua Yang, Bjarne Stig Andersen, Mats Dahlbom, Shiyu Zhuang



Load status of the operational Cray XT5 cluster at DMI in Oct 2012, prior to the implementation of DKA37, when HARMONIE forecast of 36h was made 4 times a day



Load status of the operational Cray XT5 cluster during Jan 2013 with implementation of DKA37 which generates 2-day HARMONIE forecast with 3h update frequency

Copenhagen 2013

Danish Ministry of Climate, Energy and Building

Copenhagen 2013

DMI Technical Report

Colophon

Serial title
Technical Report 12-21

Title
DKA37: HARMONIE Upgrade in December 2012

Subtitle

Author(s)
Xiaohua Yang, Bjarne Stig Andersen, Mats Dahlbom, Shiyu Zhuang

Other contributors

Responsible institution
Danish Meteorological Institute

Language
English

Keywords
NWP, HARMONIE, Asynoptic base time, Rapid update cycle

Url
www.dmi.dk/dmi/tr12-21

ISSN

Version
10 June 2013

Website
www.dmi.dk

Content

Abstract

In December 2012, DMI-CMM (Center for Meteorological Modelling) implemented an upgrade of the HARMONIE forecast suite for Danish area. The new system is named **DKA37**, where "DKA" denotes the model domain centered around Denmark, with 2.5 km grid distance and a 3-D grid mesh of 800x600x65, and "37" refers to the cycle number of the corresponding reference HARMONIE code serie 37h1. Compared to the previous real-time suite DKA which is based on HARMONIE 36h1, the new DKA37, apart from the source code update, features a 3-hourly data assimilation with 3DVAR, and a 58 h forecast produced 8 times a day. An innovative feature in DKA37 is the configuration of data assimilation centered around *asynoptic* base time, in which the analysis and starting time for forecast integration is shifted by 1 h ahead of the "conventional" base hour. Thanks to the new design about the shifted base time, DMI is able to implement a computationally intensive HARMONIE RUC (Rapid Update Cycling) assimilation alongside the other existing operational suites, resulting in a significantly improved utilisation of the HPC (High Performance Computation) resources at DMI. Compared to the situation prior to the upgrade, the new HARMONIE mesoscale forecast is not only available twice as frequent and for longer forecast length, it is also delivered 1 h earlier than the corresponding HIRLAM-SKA forecast, and 2-h earlier than the previous HARMONIE products. From parallel studies, it has been shown that the new suite, despite of a much earlier (and more frequent) delivery, is able to maintain forecast quality, with conventional verification scores, such as those for screen level parameters of wind, temperature, precipitation and mean sea level pressure, to be comparable or better than those of the HIRLAM-SKA and ECMWF. DKA37 also improves on surface wind and cloud cover over that of DKAwhich was based on CY36h1. Moreover, DKA37-RUC is found to have improved spin-up in the forecast of moisture properties, reducing jumpiness in moisture-related parameters between consecutive cycles. Although no remote sensing data has been assimilated actively in DKA37 yet, the new 3DVAR-RUC starts to assimilate such data in passive mode, opening technical possibility to explore use of AMSU, GPS-RO, ASCAT and Radar data in the coming future. The new suite also offers a further extended output parameters for end users. Technically, DKA37 is the first operational implementation of the reference HARMONIE at DMI in which the entire system is maintained in a transparent fashion via a Subversion branch in the HIRLAM source code repository.

In this report, we present the background of the HARMONIE model upgrade and example of validation results as compared to the existing NWP suites. The report documents many implementation aspects, including deviation from the reference HARMONIE 37h1.2, operational schedule, monitoring information, encountered problems and backup considerations. Finally, we offer perspectives about near future outlooks.

Resumé

DMI-CMM (Center for Meteorological Modelling) har opgraderet HARMONIE prognosesystemet for det danske område i December 2012. Den nye HARMONIE opsætning kaldes DKA37, hvor DKA refererer til modeldomænet, som er bestemt af et 800x600x65 gitter med en horisontal gitterafstand på 2,5 km centreret omkring Danmark og 37 refererer til cyklus nummeret for den tilhørende HARMONIE referencekode, 37h1. Sammenlignet med den tidligere synoptiske opsætning, DKA, som er baseret HARMONIE 36h1, vil den nye DKA37 udføre 3DVAR dataassimilering hver 3. time samt en 58 timers prognose 8 gange i døgnet. En nyskabelse i DKA37

er at dataassimileringen er centreret omkring *asynoptiske* referencetider, således at analysen og starttidspunktet for prognoseintegrationen er flyttet en time frem i forhold til de sædvanlige referencetidspunkter, 00, 06, 12, 18 UTC.

På grund af dette nye tidsskema med forskudte referencetidspunkter er DMI i stand til at køre en beregningsintensiv HARMONIE RUC (Rapid Update Cycle) assimilering samtidig med de eksisterende operationelle kørsler, hvilket har resulteret i en betydelig forbedret udnyttelse af HPC (High Performance Computing) ressourcerne ved DMI. Sammenlignet med situationen før denne opgradering er den nye HARMONIE mesoskala prognose ikke blot opdateret dobbelt så hyppigt med en længere prognose, men også tilgængelig 1 time tidligere end den tilsvarende HIRLAM-SKA prognose og 2 timer tidligere end det tidligere HARMONIE-DKA produkt.

Fra sammenlignende parallelle kørsler kan det konkluderes at den nye opsætning, på trods af langt tidligere og hyppigere tilgængelighed er i stand til at bevare prognosekvaliteten målt på gennemsnitlige værdier af 10m vind, 2m temperatur, nedbør og middel lufttryk ved havoverfladen. På disse parametre er verifikationen sammenlignelig med eller bedre end de tilsvarende verifikationer for HIRLAM-SKA og ECMWF prognoser. DKA37 har også forbedret prognosen for overfladevind og skydække sammenlignet med DKA baseret på cyklus 36h1. Endvidere har det vist sig at DKA37-RUC har reduceret spinup tid med hensyn til prognoser for fugtighedsegenskaber, hvilket reducerer spring i fugtighedsrelaterede parametre mellem på hinanden følgende kørsler.

Selv om satelitdata og andre fjernobservationer endnu ikke assimileres aktivt i DKA37, bliver disse data assimileret passivt, hvilket, rent teknisk, åbner mulighed for at afprøve anvendelse af AMSU, GPS-RO, ASCAT og radardata i fremtiden. Den nye opsætning giver også mulighed for flere output parametre til slutbrugerne.

Set fra det tekniske synspunkt, er DKA37 den første implementering af reference HARMONIE ved DMI for hvilken hele systemet vedligeholdes i en gren i HIRLAM’s subversion kildekode repository

I denne rapport præsenterer vi baggrunden for HARMONIE opgraderingen samt vurderer resultaterne sammenlignet med andre NWP opsætninger. Rapporten beskriver også hvordan den lokale implementering, DKA37 afviger fra reference HARMONIE 37h1 samt forskellige aspekter vedrørende resultater, overvågning, backup og fremtidsperspektiver.

1. Introduction

1.1 Challenges in operationalisation of mesoscale HARMONIE system

Since several years ago DMI has maintained real time run of the HARMONIE forecast system. The suite covering Denmark was initially of experimental nature and run on a domain of 384x400x65 with 2.5 km grid spacing. In early 2012, this was upgraded to a larger domain, DKA, with a grid mesh of 800x600x65, using the reference HARMONIE release of 36h1.4. The DKA suite runs 4 times a day, producing forecast of up to 36 h at each delivery. Data assimilation in DKA is limited to those of surface properties, while for the upper air, a so-called blending scheme is used, in which the short range forecast of the lateral boundary “host”, ECMWF model, is “blended in” for selected parameters representing large scale flow characteristics, such as surface pressure and upper air parameters of wind, temperature and humidity. However, due to limitation of available computational resources which in addition has to be shared with other operational NWP suites, the DKA-forecast production, as a computationally intensive, but non-operational, forecast, had to be scheduled after main operational HIRLAM deterministic and ensemble forecasts, resulting in a much delayed delivery (for one to two hours) as compared to other time critical operational forecasts. Figure 1 shows a snapshot of usage of computer nodes in the operational DMI Cray XT5 cluster. The figure, which has also been quoted in the report by Yang (2012a), illustrates the recent situation at DMI with use of HPC by both the operational HIRLAM and real time HARMONIE suites. The green bar in the plot shows the nodes that are active and the light blue the idle. The figure shows a regular 6-hourly pattern, in which the facility is nearly idle for roughly half of the time, while the other half is more loaded with a double peak structure that follows each other. In total, only a ⅓ of cluster capacity is in use. During the heavy load period as shown in Fig 1, the first peak (HIRLAM-suites) starts approximately 1h45m after synoptic time (i.e, 0745, 1345, 1945 and 0145 UTC). About one and half hour later the second peak load (with HARMONIE-suites) starts. After this, the cluster returns to low load. Such a phenomena with a heavily unbalanced load on HPC is likely common in NWP centers, in which main forecast production is organised around synoptic time. Especially so for those maintaining several time critical NWP system using similar base hour and cutoff time for input data. Such a configuration around synoptic time makes the un-balanced HPC load a structural problem. which is manifested by a massive load period with different forecast system competing on resources, followed by a strikingly idled system for rest of the time.

Ironically, a mesoscale forecast system such as HARMONIE has in general a stronger need for faster and frequent delivery than coarser resolution, synoptic scale NWP. This is due to the fact that, the main potential of the HARMONIE forecast system lies in its ability to forecast high impact weather, which is often associated with smaller temporal and spatial scales. Clearly, an earlier and more frequent delivery is essential to operationalise HARMONIE

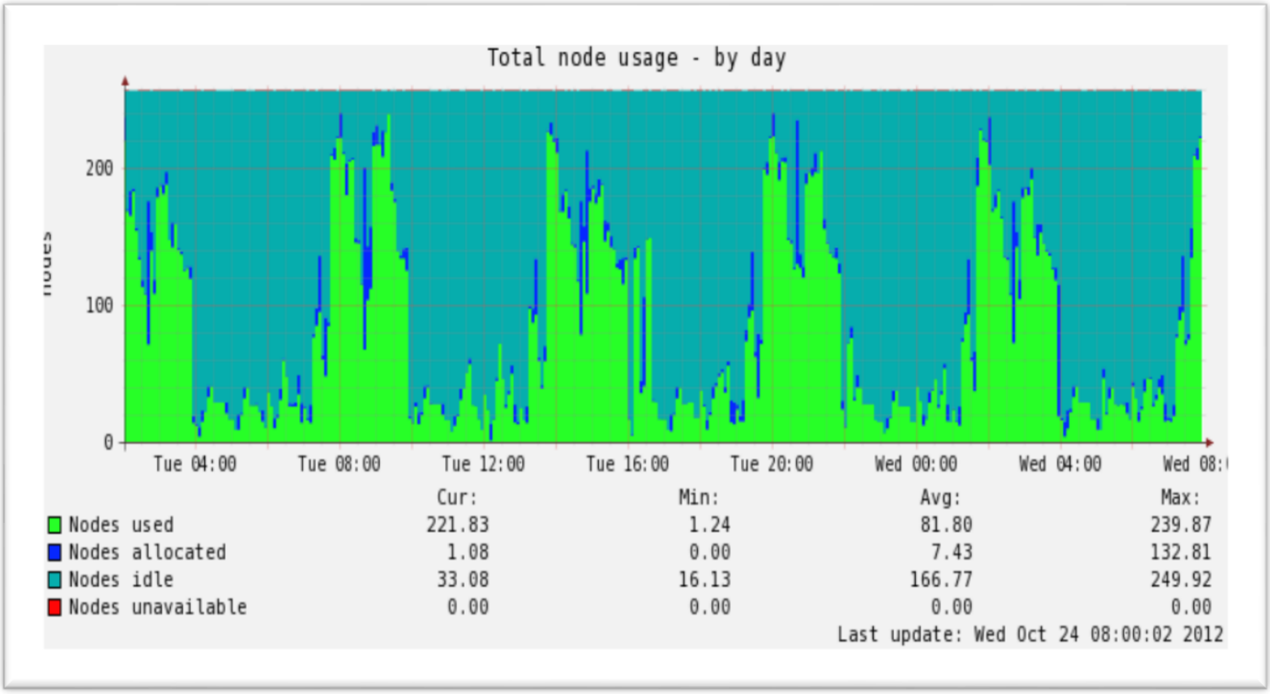


Figure 1: Example of load situation on DMI's operational Cray XT5 cluster ("Hugin") during 23 and 24 Oct 2012. (Source: DMI internal monitoring web site dmirrd.dmi.dk)

1.2 Configuring Harmonie data assimilation with asynoptic base time

In view of the above challenges in the operational implementation of HARMONIE system, Yang (2012a) conducted a feasibility study in which an HARMONIE data assimilation cycling with asynoptic base time is examined. The study consists of comparative numerical studies, in which the daily 6-hourly cycled DKA-suite as run at the DMI HPC is compared to a similarly configured one at ECMWF HPC but with asynoptic base time. In both of these suites, upper air blending scheme is used, with data assimilation applied only for surface. In the second experiment, a 3-hourly 3DVAR suite using large scale background blending scheme (LSMIXBC-option) is configured. Both of the latter experiments have been run for over two months. In these experiments, base time of the cycling is shifted 1 hour ahead of the usual, "synoptic" time. Accordingly, the assimilation data window, which is centered around the base time, is shifted 1 h ahead. The main motivation for such a shift in base time is to make it possible to move ahead usual launch time, such that the operational HARMONIE suite can be placed in the light load period on the HPC facility of the operational services, which not only utilise HPC facility better and more efficient, but also enables a more timely delivery of HARMONIE forecast. In addition, it is considered that experience with such configured data assimilation cycling is useful for further development of the HARMONIE data assimilation into the kind of Rapid Update Cycle (RUC) system, in which more focus is shifted toward assimilation of local and high resolution observation data and at further reduced time interval. As analysed in Yang (2012a), the main potential drawback of a cycling with asynoptic base time is the absence of the majority of the radiosonde (TEMP) observation data in data assimilation, due to a much earlier data cutoff in such configuration. Traditionally, TEMP data has been shown to be one of the backbones of the global observation network for synoptic scale NWP, especially in those without remote sensing data. In the case of HARMONIE 3DVAR RUC, however, recent studies have shown strong indication that the overall dominance of such data has been significantly reduced (Brousseau et al, 2012). In such a system, the limited domain size, increased cycle frequency and a short data cutoff will anyhow reduce the amount of usable TEMP data which requires longer delivery time. Meanwhile, by applying large scale background mixing (LSMIXBC) on top of 3DVAR as implemented by Vignes O. (2012), in which the short range forecast of ECMWF model is “blended” with the background states of the HARMONIE model itself via spectral nudging, the impact of assimilating TEMP data becomes even less significant. Indeed, from monitoring of the experiments, mainly the intercomparison of observation verification in comparison to that of regular run of DKA, little evidence has been found which would indicate an overall negative impact in an asynoptic-based DKA37 suite in comparison to that with synoptic basetime, despite of the fact that the former is launched and delivered significantly earlier. Interestingly, it was also found from the same study, through comparison of forecast diagnosis, that the 3-hourly 3DVAR suite appears to have a clear benefit for moisture spin-up for the short range forecast than those with 6-hourly upper air blending, indicating an encouraging potential with a 3DVAR-RUC system using HARMONIE physics.

1.3 Harmonie 37h1

In June 2012, HIRLAM-B programme released the official HARMONIE forecast system based on Cy37, 37h1.1. On Dec 20 2012, the second update, 37h1.2, is released. 37h1 code serie is a HIRLAM adaptation to the Meteo France system 37T1 in terms of Fortran source, accompanied with HARMONIE system scripts and utilities. In comparison to 36h1, 37h1 features numerous new, improved features, among those relevant to DKA37:

forecast model:

- upgraded surface scheme with SURFEX v6.1.
- updated condensation scheme EDMFM for shallow convection in AROME physics.
- improved surface spin-up due to correction on initial conversion of frozen soil from external model.
- corrected inconsistent treatment of land sea mask between boundary data and model and in climate generation.
- reduced forecast noise by using native start file from own analysis as first boundary.
- diagnosis of cloud height according to WMO definition.

assimilation:

- large scale background mixing with ECMWF forecast (LSMIXBC).
- use deep soil temperature as proxy for lake temperatures.
- optional radar data assimilation module including those of pre-processing.
- correction in VARBC module for atovs data.
- tuning about coupling between surface analysis increments of T2M and RH2M and those of deep soil temperature.

technical:

- open-mp adaptation and other i/o.
- correction of several out of bounds and portability bugs.
- improvement and update on scripts and mini-SMS scheduler.

verification and diagnostics:

- new and corrected skill scores.

- monitoring of time evolution of gridpoint and spectral norm, CPU consumption, averaged surface assimilation increment.
- several bug-fixes in FA/LFI to GRIB conversion and new grib conversion table.
- extended list of post-processing.

Prior to the release of 37h1.1, an intensive technical and meteorological evaluation was organised within HIRLAM-B programme, involving researchers from all of the HIRLAM member institutes (Yang, 2012b). The validation included examination of forecast behaviour of the new model version in different conditions, using configurations relevant to the operational community. The process also triggered several important iterations about the code and scripts, many of which resulted in further corrections and update. Since then, 37h1.1 has been further corrected and extended involving different components, such that a second tagging of the version, 37h1.2, is made on Dec 20 2012. Quality assurance for 37h1.2 involved validation study for selected winter and summer episodes and on area centered around Denmark, which showed an overall slightly positive impact (Ulf Andrae and Xiaohua Yang, personal communication).

1.4 Configuration of DKA37

Taking the aforementioned factors into account, DKA37, an implementation of the recent reference HARMONIE 37h1.2 on DKA domain with application of asynoptic-time based 3DVAR-RUC data assimilation, has been configured and tested on the DMI HPC platform, Cray XT5. Compared to the earlier 36h1.4-based suite DKA, the new upgrade has following characteristics:

- it is based on the latest released HARMONIE version, harmonie-37h1.2.
- it is a full data assimilation suite with upper air 3DVAR as well as surface assimilation.
- it is a RUC system, with a 3-hour cycling.
- it is configured with a cycling structure applying asynoptic base time, which enables a 2-h earlier delivery of forecast products compared to the situation before.
- it delivers 58h forecast 8 times a day which can be used for regular 2-day forecast, and updates with 3-h frequency.
- it starts to assimilate remote sensing data at passive mode. This means that remote sensing observation, some of them new types and have potentially high spatial and temporal resolution, are assimilated and monitored during regular runs, even though they do not yet affect final results. A passive monitoring of remote sensing data is an effective and necessary step toward the full use of data in the future.
- it post-processes additional forecast parameters, such as fog, wind gust, maximum and minimum temperature, visibility, cloud base, cloud top, icing index, etc.
- it is version-controlled under the framework of Subversion source code management, through a special code branch within the HIRLAM repository.
- it is scheduled together with other forecast suites in the HPC resource reservation system so as to ensure sufficient computation capacity.

After intensive preparation and parallel tests, DKA37 was judged ready for quasi-operational runs at the end of 2012. After further fine tuning and integration into the DMI production system, it was integrated in the NWP production system in April 2013.

In this report, we document the configuration, preparation and validation concerning DKA37 as implemented at DMI. The report also details some technical aspects related to this upgrade.

Finally, we give some summary remarks including encountered and known deficiencies, as well as perspectives on continuing efforts in the near future.

2. Pre-operational evaluation

As summarised in Chapter 1, DKA37 distinguishes from its predecessor (DKA) on numerous aspects. For most of the meteorological deviations included in DKA37, separate impact study, and in many cases documentations, already exist within the HIRLAM-B community, such as those about evaluation of 37h1.1 (Yang, 2012b), 37h1.2 (Andrae and Yang, personal communications), 3DAR with LSMIXBC option (Vignes, O., 2012), 3h vs 6h cycling (Lindskog & Thorstensson 2011), and assimilation cycling with asynoptic base time (Yang, 2012a). In this report, we limit the review about the evaluation of meteorological aspects for DKA37 to results obtained since the setup of real time runs of DKA37 from Nov 2012. Real time running of DKA37 and its evaluation is a crucial step in the quality assessment and assurance of the HARMONIE forecast system for operational use.

2.1 Pre- and quasi-operational configuration

Based on the feasibility study at the ECMWF platform with a 3-hourly HARMONIE 3DVAR cycling centered around asynoptic base time (Yang 2012a), taking into account the HPC resource at DMI and local reception time of the 6-hourly cycled ECMWF forecast as used for lateral boundary condition (LBC) in the HARMONIE coupling, a “pre-operational” DKA37 suite was configured on DMI Cray XT5 and in real time run since beginning of Nov 2012. DKA37 was based on source code of the latest HARMONIE 37h1-trunk. After several iteration involving both source and script updates, the suite was upgraded to that of 37h1.2 and the tagging based on branch of DMI/harmonie-37h1_oper on Dec 20 2012 became the system base since then. In Jan 2013, the configuration went through further changes to enter the “quasi-operational” phase. During this stage, forecast from DKA37 starts to get integrated into production system and accessible by duty forecasters. But as an non-operational suite, the duty of maintenance is still on developers side.

During the pre-operational phase, DKA37 was configured with an uneven cycle, which consists of “production cycle” with normal forecast lead time of 37h and “intermediate cycle” configured for data assimilation purpose. For intermediate cycles, only a 3h forecast is run in order to provide background model state for next cycle. The production cycles are specified for 23,05,11,17, providing HARMONIE forecasts matching those of the other NWP forecasts with base time of 00, 06, 12 and 18. With this implementation, the launch time for shorter cycles is adjusted so as to take advantage of both the late arrived observation data and the LBC from the latest ECMWF cycle. In Table 1, the daily schedule as implemented at DMI in Dec 2012 is shown, in which special arrangement has been made for short cycles.

The delivery time of 37h forecast from DKA37h as shown in Table 1, demonstrates one of the most remarkable advantages with DKA37 in comparison to the other available routine forecast products at DMI. Before DKA37, e.g., for 12 UTC cycle, the 36h forecast of HIRLAM-SKA is available at about 14:30. For DKA, it is around 1600. With the new DKA37, the equivalent forecast for same valid time is around 13:30. Thus, DKA37 forecast is delivered 1 h earlier than equivalent HIRLAM forecast, and 2h30m earlier than the earlier DKA forecast.

Table 1: DKA37 schedule during "pre-operational" phase (Dec 2012)

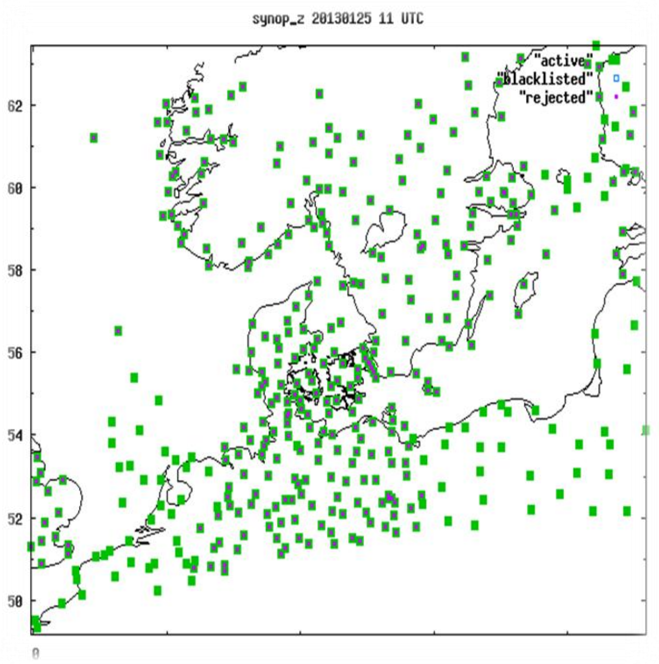
Base hour	Assimilation window	Obs cutoff	Cycle origin of ECMWF LBC	Launch time	Forecast length	Estimated output time of final forecast
02	00:30 – 03:30	05:55	00	05:55	3h	-
05	03:30 – 06:30	06:15	00	-	37h	07:30
08	06:30 – 09:30	11:55	06	11:55	3h	-
11	09:30 – 12:30	12:15	06	-	37h	13:30
14	12:30 – 15:30	17:55	12	17:55	3h	-
17	15:30 – 18:30	18:15	12	-	37h	19:30
20	18:30 - 21:30	23:55	18	23:55	3h	-
23	21:30 – 00:30	00:15	18	-	37h	01:30

For operational implementation, starting from "quasi-operational" stage, the above schedule is further modified, in which the cycle configuration changes to an even-cycled structure with 58 h forecast for all 8 cycles. This is detailed in Chapter 3.

2.2 Use of observations

Conventional observation data used in data assimilation of DKA37 includes the pressure observation from surface SYNOP, SHIP, DRIBU, and upper air observation with radiosonde (TEMP), aircraft and airep (AIREP).

Configuration of 3DVAR RUC with asynoptic base time in DKA37 implies an observation data cutoff time which is significantly earlier than for the case typical to regular NWP suites at DMI and other NWP centers. Figure 2 shows observation data distribution maps for SYNOP, SHIP, aircraft and radiosonde data, respectively, for cycle 11 UTC, Jan 25, 2013, in which an observation data cut-off of 1h15m after base time, i.e., at 12:15 UTC, is applied. As shown in the figure, due to a rather early data cutoff that is very close to synoptic time, only very few TEMP data is found in the DKA37. On the other hand, for other in-situ observation data, thanks to the fact that most of such observation are nowadays automatic and continuous, with multiple report per hour and typically with little delay, there appears to be little influence on the amount of available observation data as consequence of a shorter data cutoff. One exception of this is SYNOP data from some East European countries such as Poland, Baltic states, etc., which is available to DMI only at 3h frequency. Thus, for some of the DKA37 cycles, the available observation data depends strongly on selection of observation data cutoff time. Figure 3 illustrates this with a SYNOP data usage plot which is similar to Figure 2a, but for a test cycle 2013011911, in which the observation data cutoff time is moved ahead by 5 minutes to +1h10m. It is shown that this 5-min earlier cutoff resulted in a significant loss of SYNOP data, mainly on the south-east corner of the domain.



.
Figure 2a.

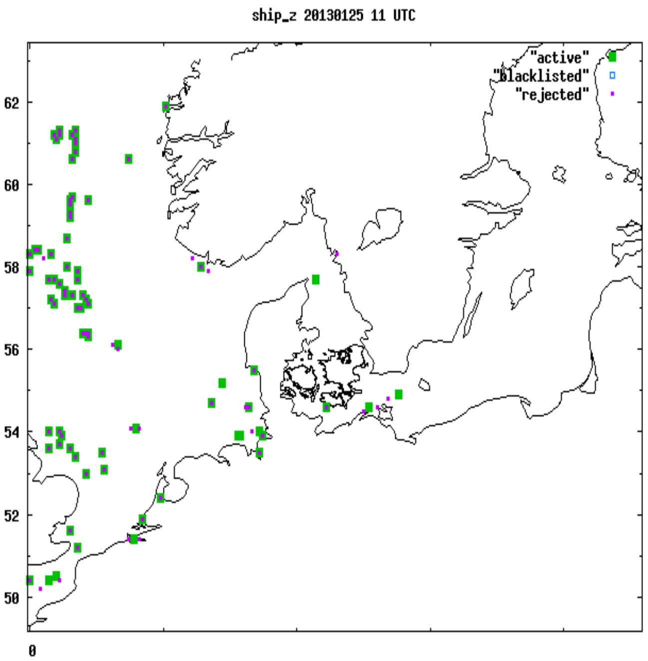


Figure 2b.

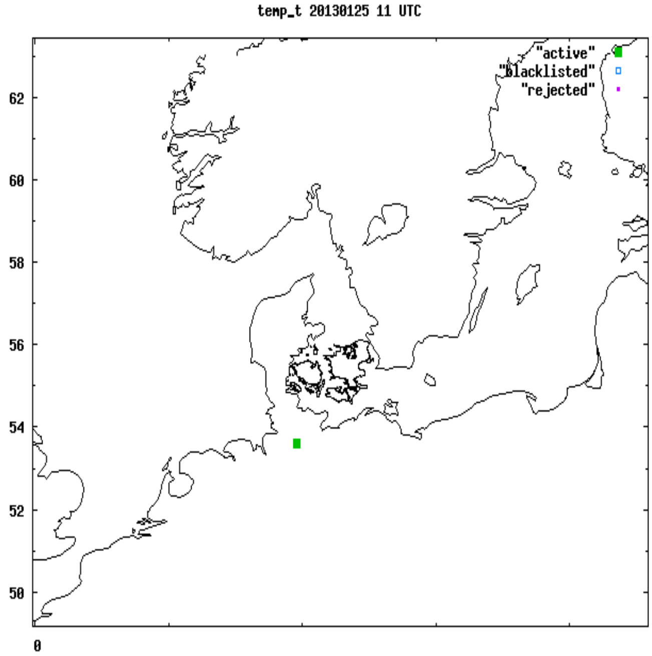


Figure 2c.

Figure 2. Actively assimilated observation from SYNOP stations for cycle 2013012511, with cutoff at 12:15 UTC, with a) SYNOP surface pressure data; b) SHIP and DRIft BUoy (DRIBU) surface pressure data; c) Aircraft temperature and wind speed profiles; d) radiosonde (TEMP) data.

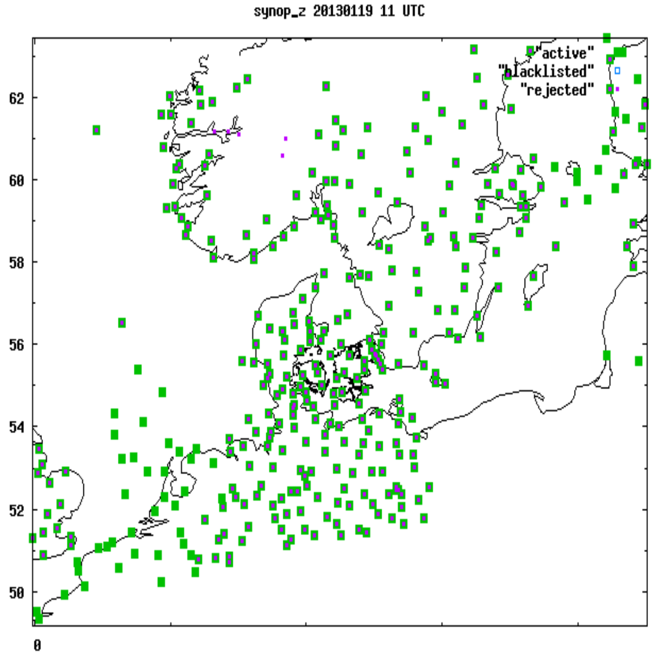


Figure 3. Actively assimilated surface pressure observation from SYNOP stations for cycle 2013011911, with cutoff at 12:10 UTC.

In DKA37, three types of remote sensing data: ATOVS radiances, Ascat wind and GPS-radio occultation data, start to be included in the input observation data in passive assimilation mode. These are to be monitored regularly, with a view for activating in the near future. The ATOVs observations currently available for use in DKA37 comprises of radiance data from the AMSU-a and AMSU-b/MHS instruments onboard the NOAA-15, 18, 19 and Metop-A satellites.

For radiance data, the algorithm of variational bias correction is targeted for active assimilation. The recent modifications in HARMONIE-37h1.2 as recommended by Lindskog et al. (2012) have been implemented into DKA37, which includes exclusion of three predictors and a tuning of the predictors scaling factors.

The advanced scatterometer instrument is situated onboard the Metop-A satellite and produces sea wind data with a 12.5 km resolution.

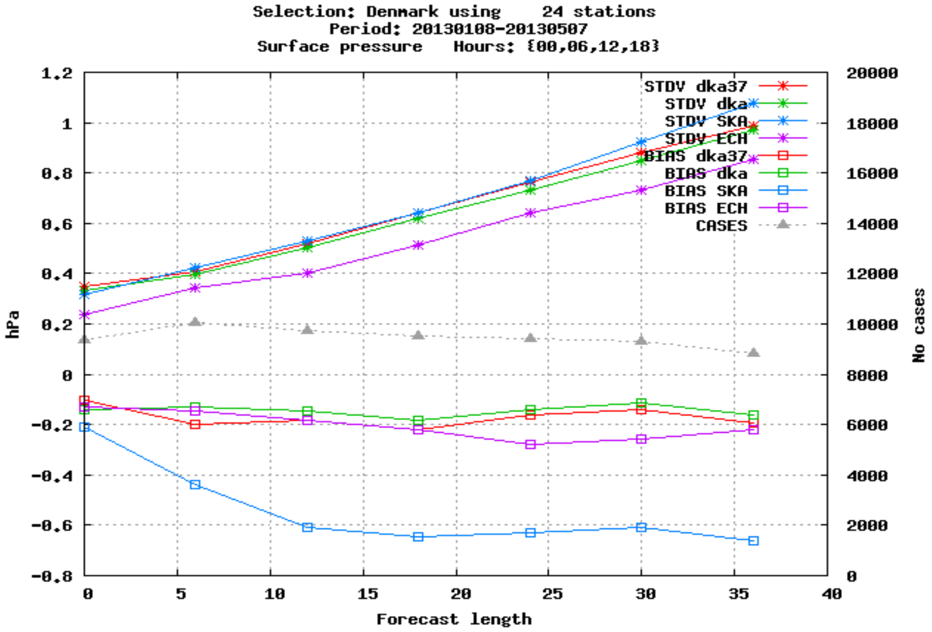
GPS radio occultation data are assimilated as bending angle data that are converted to humidity profiles. The data originates from several different SAF projects, including GRASSAF, COSMIC, GRACE, CHAMP and TerraSAR-X. These observations have a very high vertical resolution, hence may potentially compensate, to some extent, the reduction or absence of radiosond data. Furthermore, since these data contain complementary information with respect to the ATOVS data, it has been found in global NWP analyses that this observation type can act as an anchor to the VarBC system.

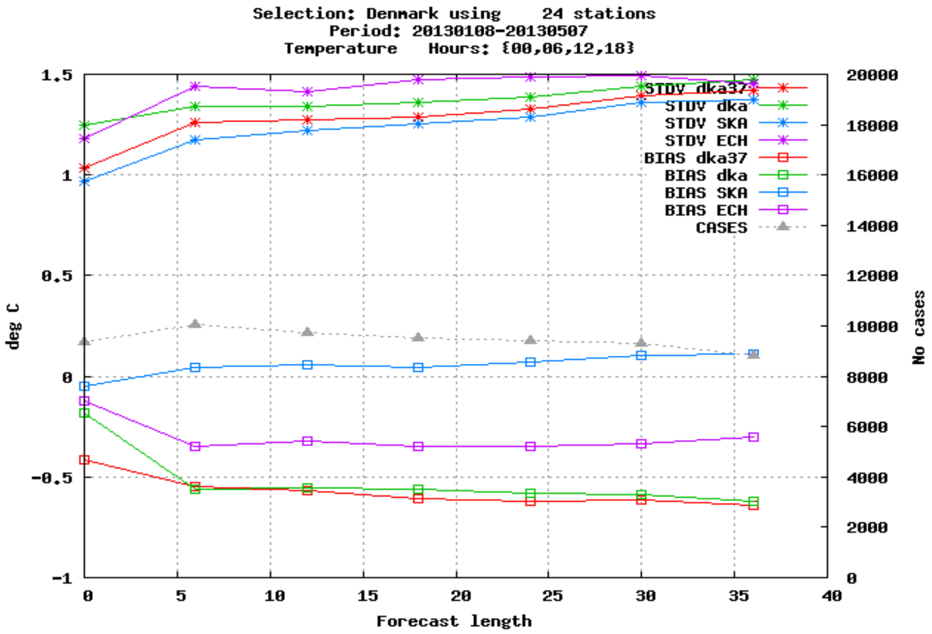
The blacklisting for the above mentioned observations has been modified so that, at present, all remote sensing data are set to passive mode and will be monitored until a stable set of coefficients has been reached for some of the data types (such as ATOVS) and a non-deteriorating impact on the forecast system has been established. In this process, experiences from assimilation of radiance data in DMI-HIRLAM has been useful. For DKA37, similar data selection about ATOVS satellites and channels used in DMI-HIRLAM has been applied.

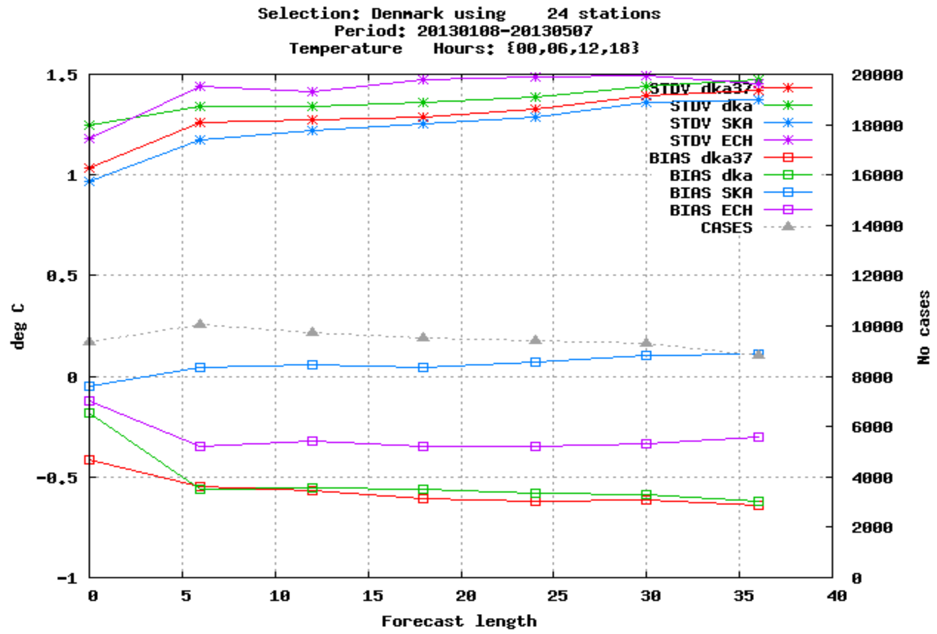
2.3 Observation verification

For observation verification, the evaluation includes intercomparison between DKA37 and the other routine forecast products available to the forecasters in DMI: HIRLAM (SKA, 3km grid distance) , ECMWF (ECH, 16 km grid distance) as well as Harmonie-DKA based on 36h1.4 (DKA). In order to enable comparison of model results with different base time, the forecast output from DKA37, which runs on asynoptic base time with 1h shift ahead of usual synoptic time, is compared to other model results with 1h shift, thus, the inter-comparison is slightly unfavorable to DKA37, as its actual forecast lead-time is 1h longer than other models.

Figure 4 (a to b) shows conventional synoptic scores measuring averaged model errors in standard deviation and bias, as compared to surface SYNOP observations, for models in comparisons, in mean sea level pressure (MSLP), 2m temperature, 10 m wind speed and total cloud cover, respectively. The statistics cover a period of 4 winter months during Nov and March 2013. It is seen that, for MSLP, the STD error by ECMWF model is, as usual, superior to those of the LAM models. Further, DKA37 suffers a marginal degradation at nominal analysis time, presumably due to 1h shift of analysis time. For the bias error, HARMONIE forecasts are comparable to that of ECMWF and superior to that of HIRLAM-SKA.


Selection: Denmark using 24 stations
Period: 20130108-20130507
Surface pressure Hours: {00,06,12,18}
STDV dka37
STDV dka
STDV SKA
STDV ECH
BIAS dka37
BIAS dka
BIAS SKA
BIAS ECH
CASES
hPa
No cases
Forecast length


Selection: Denmark using 24 stations
Period: 20130108-20130507
Temperature Hours: {00,06,12,18}
STDV dka37
STDV dka
STDV SKA
STDV ECH
BIAS dka37
BIAS dka
BIAS SKA
BIAS ECH
CASES
deg C
No cases
Forecast length



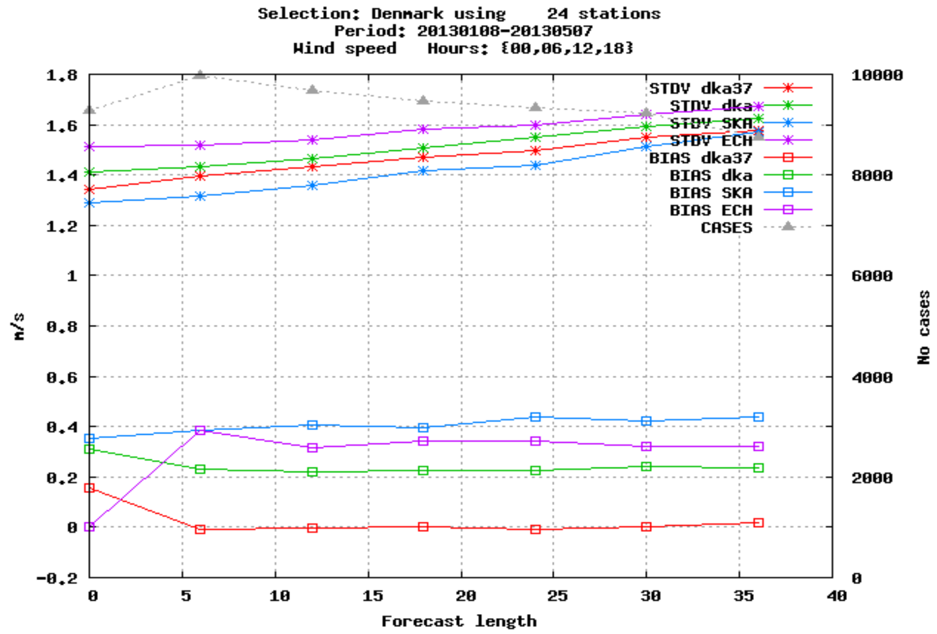
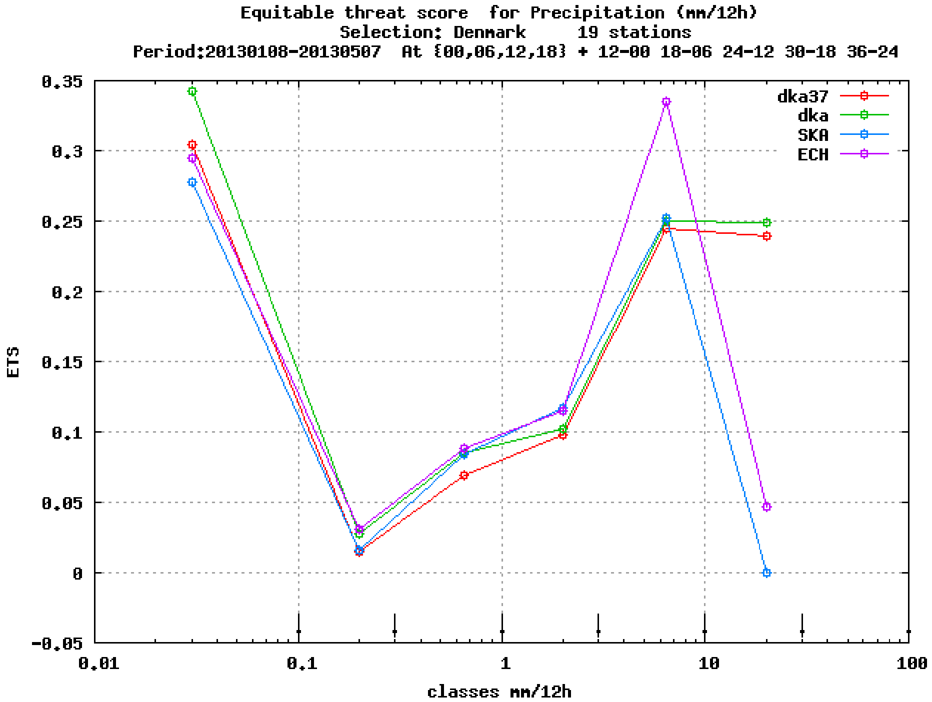


Figure 4. Observation verification using surface SYNOP observation for the full DKA domain, comparing pre-operational DKA37 (red) against DKA (green), HIRLAM-SKA (blue) and ECMWF (magenda) for Jan 8 to May 7 2013, a) MSLP; b) T2m; c) W10m, d) cloud cover. Note the dark dotted lines are for number of paired data, showing a significantly smaller value for forecast range beyond 24h. The latter is due to limited number of runs with forecast range beyond 24h during pre-operational stage

While the scores for the MSLP, traditionally considered a parameter indicative of the prognostic skill for the synoptic scale, is favorable to the ECMWF model, the STD error of T2m and W10m as shown in 4 (b) and (c) demonstrate a clear advantage of LAM models, both for HARMONIE and HIRLAM, over that of ECMWF, presumably due to resolution. For bias, HARMONIE models suffer a negative temperature bias of over half degree for Danish stations, whereas for HIRLAM SKA and ECMWF, overall the bias are less pronounced. For surface wind, the bias in both DKA37 and SKA are small, with DKA and ECMWF forecast suffering larger positive bias.

The cloud cover verification shows a larger STD error with LAM models compared to that of ECMWF. This presumably can, at least in part, be attributed to a double penalty tendency with high resolution model, especially in cloud and precipitation resolving small scale structures. The cloud bias over Danish area seems to be on positive side, but the spin-up in DKA37 is clearly better than in DKA, presumably associated with use of 3DVAR-RUC. This is consistent with the similar finding in Yang (2012a).



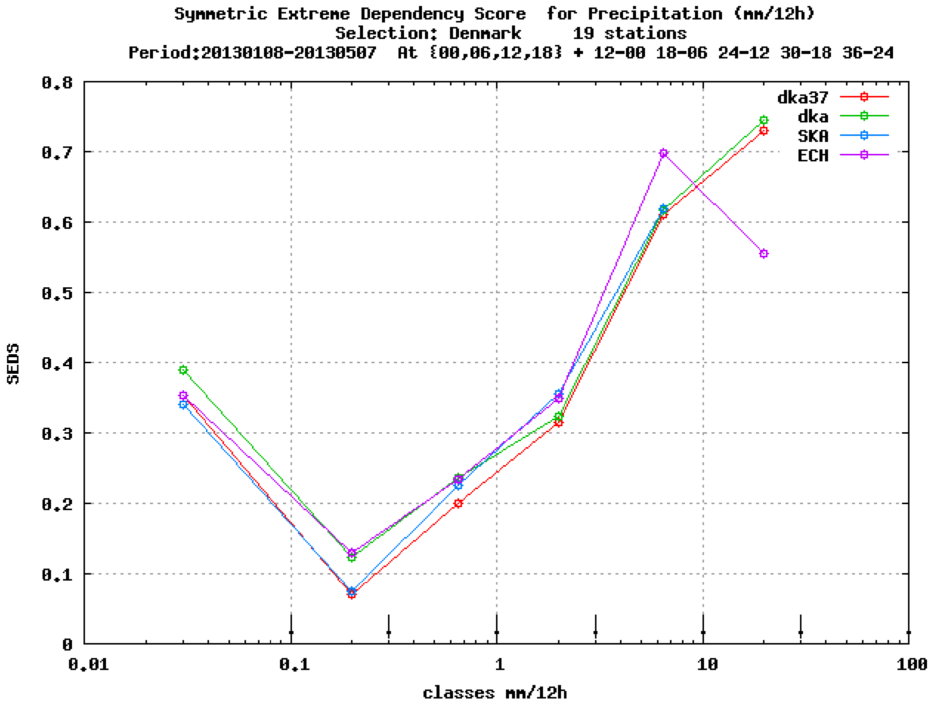
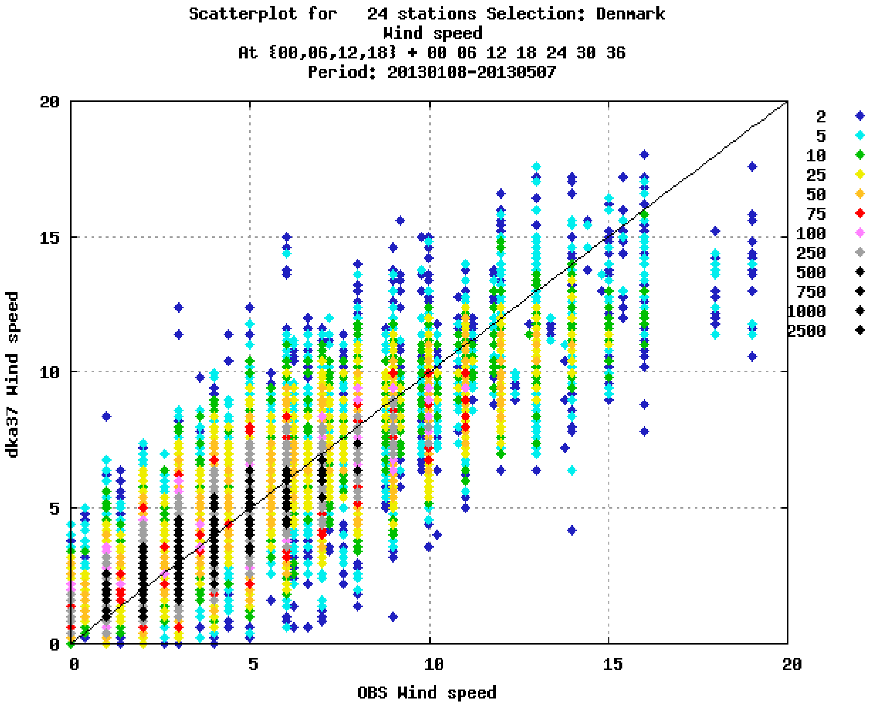


Figure 5. Same as in Figure 4 but for observation verification of 12h accumulated precipitation with Equitable Threat Score (a) and Symmetric Extreme Dependency Score (b).

Figure 5 displays two conventional scores for precipitation forecasts over Danish area, with equitable threat score in (a) and symmetric extreme dependency score in (b) for different thresh-hold. Interestingly, the coarse resolution model ECMWF shows an advantage in the intermediate threshhold, whereas for HARMONIE, the main advantage is shown in the high precipitation threshhold. Despite of usual deficiency with conventional point verification for verification of precipitation events with high resolution model, it is interesting to see that even with conventional score, HARMONIE is still able to demonstrate some added values in forecast of high magnitude precipitation events.

Figure 6 and 7 displays scatter-plots for observation verification of 10-m wind over scandinavia area and for mountain list, respectively, comparing models of HARMONIE-DKA37 (upper) vs those of ECMWF (lower). The comparison seen from these plots indicate that, while DKA37 has in general a less scattering (hence standard deviation error) of predicted wind speed than that of ECMWF forecast in relation to the observed wind speed, both models suffer clearly underprediction of wind in case of strong wind above ca 15 m/s. This is especially the case for mountain list, for which there appear to be a general underprediction of wind speed especially for high wind cases. For screen level temperature as shown in Figure 8, both of the models look to have difficulties in predicting cold winter time temperature (< -30 degree) in the nordic area. Such is especially severe for the ECMWF forecast. A deeper investigation of the cases with over prediction of wintertime surface temperature shows that these are often associated with situation with clear and calm conditions (not shown here), probably indicating a general model deficiency in representing flow situation with a stable boundary layer.



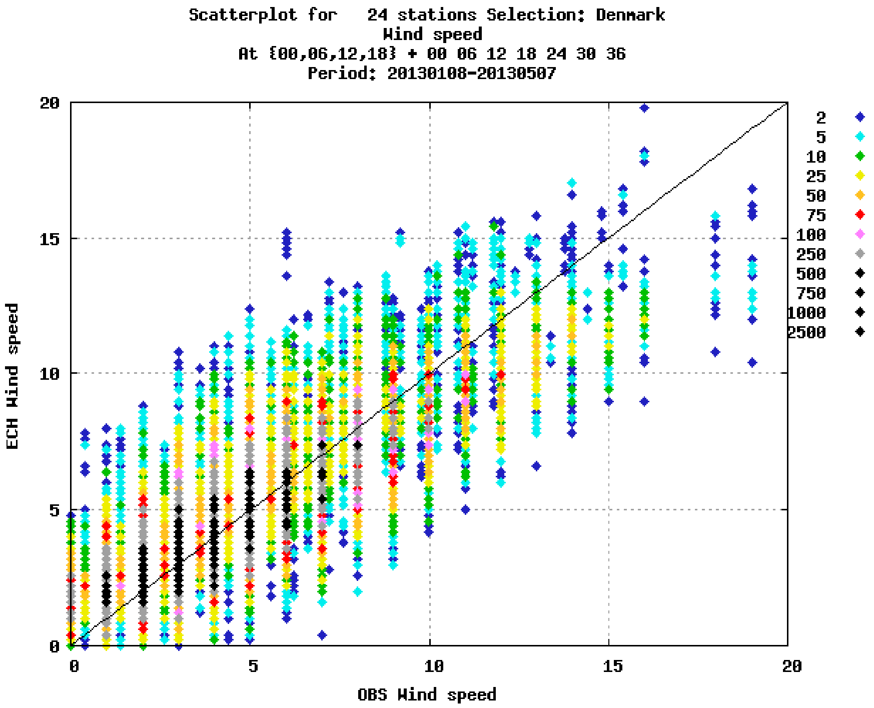
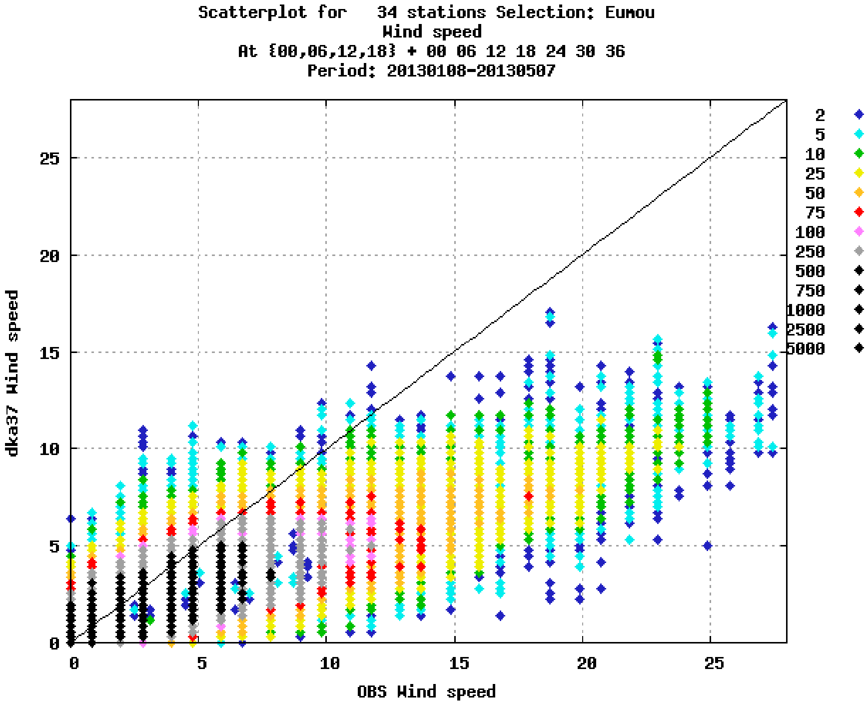


Figure 6. Scatter plot showing observation verification of surface wind observation for scandinavia station list, comparing DKA37 (upper) and ECMWF (lower) for Jan 8 to May 8 2013. The X-axis for observation data and Y-axis for model data with forecast length of up to 36h. The colours indicate number of cases



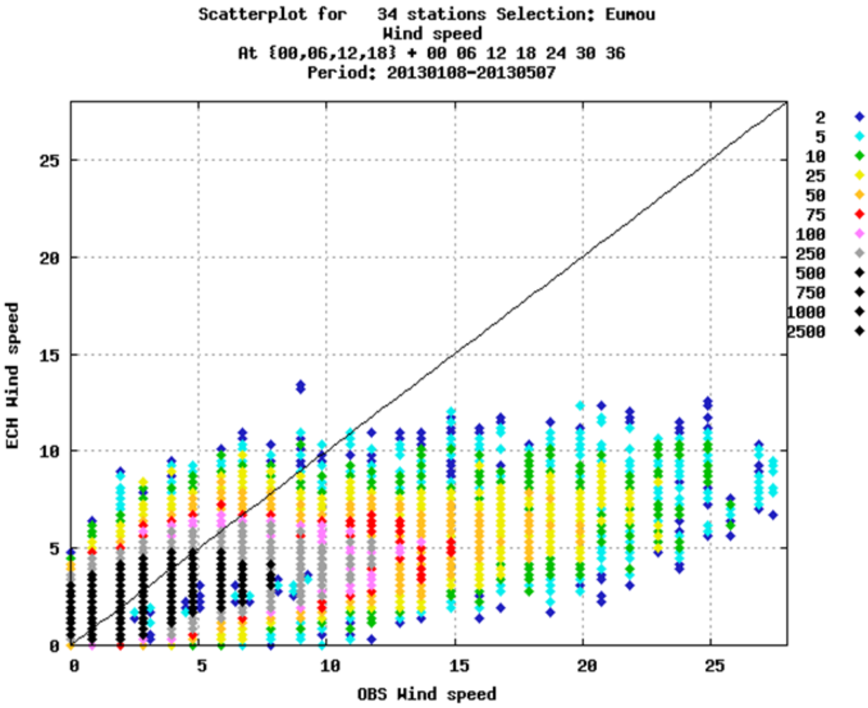
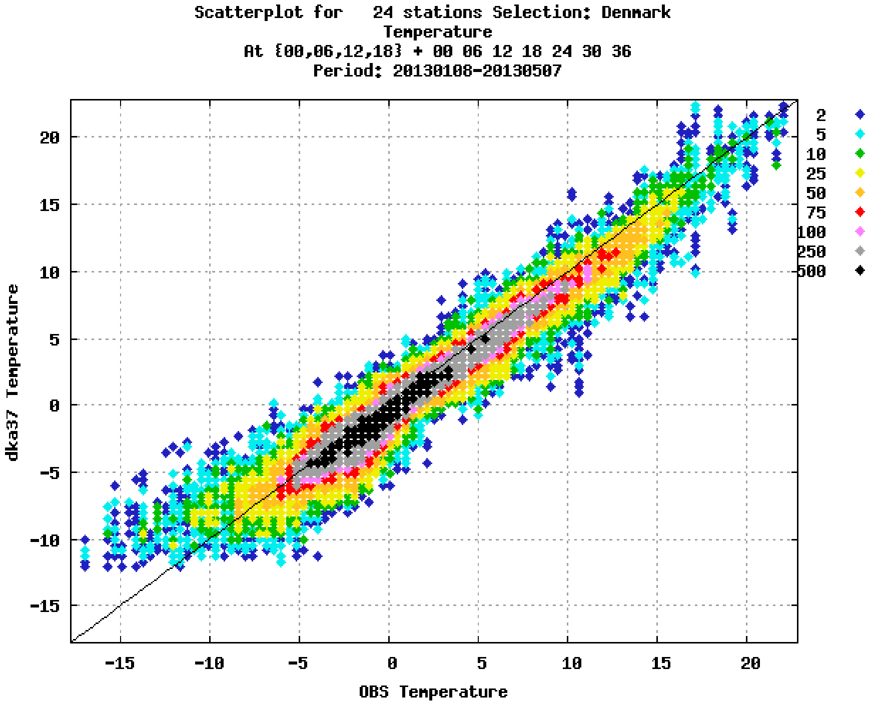


Figure 7. same as Figure 6 but for mountain station list.



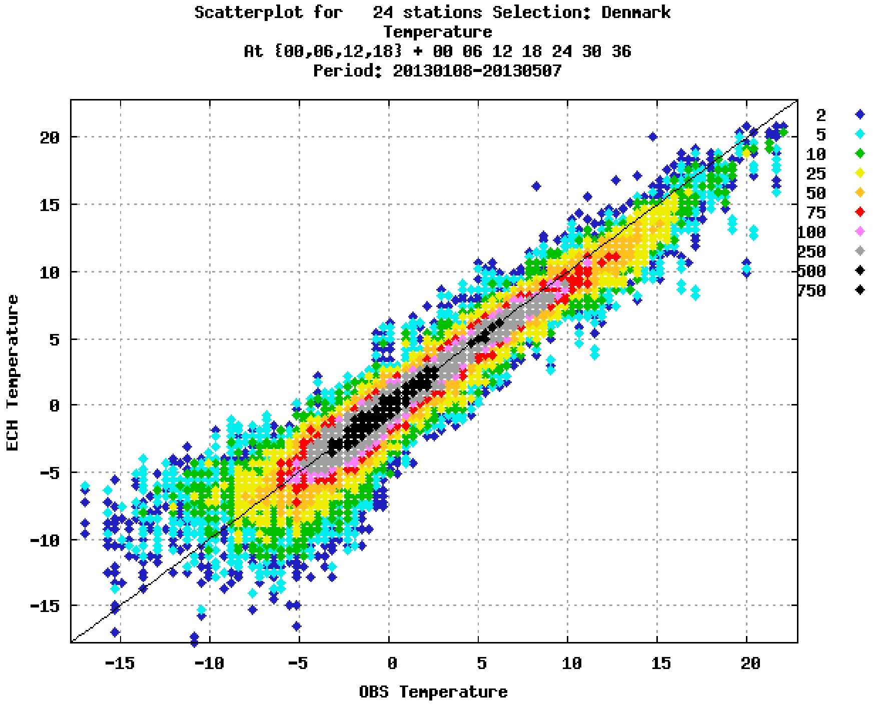


Figure 8. same as Figure 6 but for verification of temperature at 2 m.

To summarise, verification inter-comparison for conventional weather parameters between those of the Harmonie forecasts and other model results available to DMI end users, (HARMONIE-DKA, HIRLAM-SKA, ECMWF), indicates an overall competitive quality with the upgraded HARMONIE suite. Among these, DKA37 is seen to be associated with a very competitive model error in surface wind, bias error in MSLP, and standard deviation error on T2m, and an advantage in precipitation for high thresh-hold. Among the main observed deficiencies during the past winter period, DKA37 tends to have a relatively large negative T2m bias, a tendency of overprediction of T2m during clear and calm winter condition, an underprediction of strong wind, especially too weak over mountain area during high wind situation. For cloud cover, the DKA37 scores improve those of DKA but still suffers larger standard deviation error than in HIRLAM-SKA and in ECMWF. But all in all, the overall meteorological performance of DKA37 is satisfactory.

3. Operational implementation and experiences

3.1 Configuration of 3-hourly 3DVAR data assimilation with asynoptic base time

DKA37 has been shown to provide not only a more frequent and longer AROME forecast, but also with a significantly earlier delivery than the situation before. Table 2 lists some of the operational characteristics of DKA37 in form of daily schedule, featuring a launch time of +1h15m after each base hour, using an observation data cutoff of 1h15m. With a 60 s time-stepping and hourly output, 124 Cray XT5 nodes, the 2-day DKA37 forecast is finished within 1h30m, so that the forecast output in GRIB format can be delivered to end users well within 2h after each "regular" time (00, 03, 06... 21 UTC), with a fresh forecast every 3h.

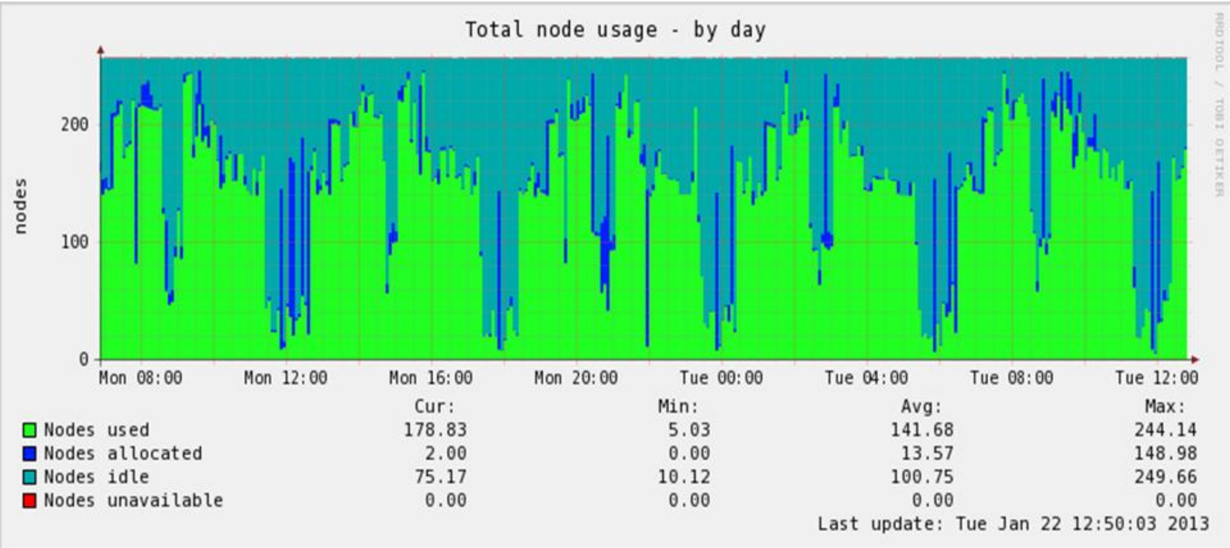


Figure 7: Example of load situation on DMI's operational Cray XT5 cluster (Hugin) in late Jan 2013, during the second phase of DKA37. (Source: DMI internal monitoring web site dmirrd.dmi.dk)

Table 2: Operational schedule for 3-hourly cycled DKA37 at DMI

(all time given in UTC)

Base time	Nominal cycle	ObsData window	Obs cutoff	EC BC cycle	Launch	Forecast length	Delivery time
02	03	00:30 – 03:30	03:15	18	03:15	58h	04:45
05	06	03:30 – 06:30	06:15	00	06:15	58h	07:45
08	09	06:30 – 09:30	09:15	00	09:15	58h	10:45
11	12	09:30 – 12:30	12:15	06	12:15	58h	13:45
14	15	12:30 – 15:30	15:15	06	15:15	58h	16:45
17	18	15:30 – 18:30	18:15	12	18:15	58h	19:45
20	21	18:30 - 21:30	21:15	12	21:15	58h	22:45
23	00	21:30 – 00:30	00:15	18	00:15	58h	02:45

Figure 7 shows example of load situation with regular DKA37. Comparing this to the previous situation as shown in Fig 1, it is clear that the new configuration improves substantially the utilisation of the HPC resources at DMI.

Cycling using asynoptic base time is new. To ease local adaptation and for user convenience, the GRIB output and other post-processed model results from DKA37 are renamed, so that the final forecast are still presented to users as those from usual cycles with base time at “regular” hour. *i.e.*, a +1-h forecast started at 05 cycle is renamed as a forecast started from ‘nominal cycle’ of 06 with lead time of 0 h. Likewise, a 37-h forecast from 05 cycle is presented to users as 36-h forecast from 06 cycle. This means that, in the DMI GRIB data base (under the directory /gdbopr in the HPC), the DKA37 output are presented in the same fashion as those from other models, *e.g.,* the data 2012121212/012 for DKA37 is in reality 2012121211/013. Such a renaming procedure has an additional advantage to enable a convenient comparison between DKA37 and other model forecasts originated from synoptic base time that is 1-h later.

3.2 Source repository

System-wise, DKA37 is the first operational NWP suite at DMI for which the complete source, including those of scripts, is maintained by a revision control system. DKA37 is an implementation of the HARMONIE version **dmi-harmonie-37h1.2**, dated Dec 20 2012, which in turn is tagged from the branch DMI/harmonie-37h1_oper within the HIRLAM source code repository. The branch spins off from the HARMONIE trunk at revision number 11169, dated Dec 5 2012, which at the time of tagging has been updated to the release version of harmonie-37h1.2 in HIRLAM-B, dated Dec 20 2012. On May 24 2013, a new tagging, dmi-harmonie-27h1.2__oper was made to reflect minor upgrades since last tagging.

On DMI Cray XT5, the source for DKA37 has been built with pathscale and intel compilers using Makeup with optimisation level -O2. Experimentation of the executibles for forecast step showed significant efficiency difference between the two, with the one with intel compiler being clearly more efficient. Fortunately, the forecast results using two executibles, while different, appear to be insignificant meteorologically speaking. Thus the executibles made by intel has been selected. Main components of the data assimilation and forecast steps are run with MPI parallelisation.

Deviation from the reference HARMONIE 37h1

Deviation of DKA37 from the reference HARMONIE 37h1.2 consists of modifications and adaptation for both Fortran source codes and scripts, but mostly on scripts side and they are mainly of technical nature. The deviations from the reference version are of the following kinds.

- **Platform adaptation** for local HPC architecture and compiler options.

- **Adaptation for local production environment** concerning handling of disseminated ECMWF LBC data, observation data format and bufr table, job schedule and reservation scripts specific to the DMI HPC, local output and post-processing data stream, GRIB data conversion, archive and disk cleaning, diverse backup measures, and launch script triggered via crontab. Among these, the handling of incoming ECMWF boundary data on the fly is a major deviation from the reference scenario. This is because that, contrary to the practice at the other operational HARMONIE implementation, as shown in Table 2, DKA37 uses latest arriving ECMWF forecast with various lagging for different cycles. This means that the interpolation procedure is launched while the latest ECMWF LBC is still arriving. This requires a robust script to handle the situation with input data either on the fly, or with substitution in case of delay.

- Script adaptation to configure **3 hourly cycling with asynoptic base time**;

- **Entry** of remote sensing data for passive monitoring, through which observation data of ATOVS amsu-a, amsu-b/mhs, ASCAT, radio occultation data, ground-based GPS, are incorporated into the observation data input, although in passive mode;

- Other minor technical adaptation.

3.3 Data assimilation aspects

Structure functions

In order to carry out upper air 3DVAR data assimilation, structure functions (also known as background error covariances) for DKA domain is required. For DKA37, the initial structure functions have been derived from a sample data set comprising of ensemble of 6h HARMONIE forecasts downscaled from the ECMWF global ensemble data. Using the daily ECMWF ENDA (ensemble data assimilation) data set during 2010060100 and 2010063000 as initial and lateral boundary conditions, 4 Harmonie ensembles, each with 6h forecast per day, have been run. With a total of 120 (30 day x 4 ensemble) samples, the statistics of structure functions are computed using statistical program, following the procedure described in the HIRLAM system wiki (https://hirlam.org/trac/wiki/HarmonieSystemDocumentation/Structurefunctions)
It shall be noted that the selection of the current structure function maybe rudimentary and a tuning about the structure function and its scaling may be beneficial in longer run. e.g., sensitivity about seasonal dependence of structure functions derived for summer episode and for winter episode have been found in a recent impact study by Shiyu Zhuang, the result of which calls for further investigation and tuning.

3.4 Encountered problems and known deficiencies

Running several months of DKA37 revealed various deficiencies, irregularities and interruptions with the system, which also have benefited the overall robustness in operational use, including introduction of some automatic fallback solutions. For all irregularities encountered in the real time suite, it is important to document them as extensive as possible, with troubleshooting measures formulated if possible.

Cycle interruptions. To date, since start of DKA37, except for script errors associated with handling and preparation of observation data and ECMWF boundary , there have mainly been two types of cycle interruptions, both of which in forecast integration step.

- One of the most frequent crashes (in average a few cycles per month) during forecast step is associated with a log message indicating “MPICH PtlEQPoll event->ni_fail_type error (OTHER EQ handle): This is likely due to an unresponsive node on the system”. The error is non-producible since a subsequent re-launch of forecast step would always make the run through. Although the error messages appear to point to hardware problems, common to all such cases is the fact that such always happen at the i/o step. Report has been forwarded to the technical department at DMI.
- Two incidences happened during the initial test period in which forecast aborts with the error message suggesting “SMILAG TRAJECTORY OUT OF ATM/UNDERGROUND N TIMES”. Various tests, including reduction of time step, replacement of condensation parameterisation scheme EDMFM by EDKF, running hydrostatic instead of nonhydrostatic mode, running with different compilers, different number of MPI nodes and domain decompositions, all failed. In these cases, running ‘blending’ scheme instead of 3DVAR, or running with slightly different initial data made the rerun through. Such type of crashes in forecast integration is worrying but no deeper debugging has been done due to time constraint, and also due to difficulties with arranging catch-up cycles.

In order to optimise utilisation of HPC resource, DKA37 is now run under a reservation system in the operational cluster with a 2h40m time-slot and up to 140 available nodes every 3 hour, which is shared with two other operational suites. While such arrangement has the advantage to optimise usage of HPC resources, it becomes a challenge when irregularity happens with cycling. This is especially so in case that DKA37 suite suffers delay or crash. To ensure smooth and reliable production, several backup measures have been constructed, as are reviewed in the section below.

Known deficiencies. Technically, several deficiencies with DKA37 have been encountered, such has since become target of follow-up efforts in optimisation tuning and some already resulted in minor upgrades:

- Computational efficiency. The present version runs with system compiled by intel compiler and with pure MPI optimisation. From other experiences a correct implementation of hybrid optimisation with mpi-openmp, the system may potentially get some speed-up in forecast step, and may also be helpful to avoid the type of un-reproducible crash as reported above.
- Especially, output step is currently a severe bottleneck in terms of computation efficiency. Solution for this is on the way to be implemented in CY38, which may be implemented at DMI in late 2013.
- Low efficiencies on Cray XT5 platform in surface assimilation, both for OI-main option and its alternative, SODA. This step involves a relatively simple interpolation to convert results of screen level surface analysis down to soil levels for temperature and humidity parameters defined in the model surface schemes. During the test, the default SODA scheme is found to take more than 10 minutes and has therefore to be replaced by its predecessor, OI-main, but even the latter consumes around 6 minutes, which is substantially slower than in CY36h1.4 and also more than the combined computation cost of upper air (3DVAR) analysis. The problem has now been addressed through efforts by colleagues at met.no (Ole Vignes) and an open-mp enabled version of OI_MAIN has now been implemented, resulted in a significant efficiency gain so that the time spent on OI-main is now reduced to be less than 2 min.
- Post-processing of forecasted or diagnosed parameters in GRIB form, as requested by DMI forecasters, is not fully settled, with some quantities showing zero or wrong values, some are still missing.
- Storage and archive strategy for DKA37 remain to be implemented. The full model state with hourly output requires extensive storage, which poses substantial burden to the HPC system.

3.5 Automatic measures to ensure regular production

Automatic and manual steps to ensure continuous and timely delivery

Various automatic measures have been added into the configuration of DKA37, with the aim to easily maintain a continued forecast production in case of anomaly. The essential strategy behind these measures is to prioritise scheduled product delivery and to maintain cycling. Considering the constraints in staffing and in computational resources, following principles and backup measures have been implemented for the critical steps in the DKA37 data flow:

- **Separate** non-essential post-processing steps from main production suite so as to avoid interruptions of the time-critical cycling in case of problems in the non-critical parts of the data flow.
- **Secure smooth handling of observation data.** An observation data cut off time is defined, currently chosen as +1h15m after base time. If for any reasons no conventional observation within the data window is available, or if such data has been found to corrupt, no data assimilation will be performed for the cycle and the forecast is initiated with short range forecast from previous cycle.
- **Lateral boundary data assured.** A boundary data cutoff time of 1h20m is defined. If any of the expected ECMWF forecast boundary is not delivered by the cutoff time, the equivalent ECMWF forecasts from the previous cycle will be used.
- **Crash in surface analysis.** The step is skipped over and the first guess for the analysis is defined as output for this step.
- **Crash in upper air analysis.** The step is skipped over and the first guess for the upper air analysis, (usually the output from the surface analysis), is defined as output for that step.
- **Crash in forecast step.** Diagnose cause of a forecast failure could be difficult. Since forecast step accounts for major part of the total production time, following arrangements have been designed
    - in case the crash happens after the 3-h forecast has been produced, no re-run will be made. Instead, cycling is resumed at following cycle.
    - in case the 3-h forecast has not been produced, re-run the cycle using upper air blending mode ANAATMO=blending to produce 3h forecast
    - in case the above step fails, skip the cycle, and run next cycle with FC=6 to define the previous successful cycle to be first guess. This scenario has so far not occurred.

Alternative product backup

Once officially accepted as operational product, forecast from DKA37 is expected to get integrated into the production chain at DMI as bases for routine weather forecast. In case of delivery anomaly due to unforeseeable reasons, the following model and products are recommended to be used as backup products, in the following priority order:

1. **DKA37's own products from previous cycles.** In view of the required time for a full production cycle, for each delivered DKA37 forecast with a targeted forecast length of 48+ hour, the forecast lead time of DKA37 is set to be 58 h, so that no second attempt is necessary in case of cycle interruption for two consecutive cycles. Instead, the extended-range forecast from previous cycles are used to maintain production. This is in consideration of the fact that, due to the extended resource request for CPU and need for earlier and frequent delivery, it has become much more difficult to schedule, within the normal resource reservation time limit, a repeated run in case of cycle interruption.

2. **DKA(-36h1).** Prior to operationalisation of DKA37, the 36h1-based DKA has been running rather stably in real time since early 2012. The output from DKA has since been delivered to forecasters for evaluation. Thus, it is natural to use DKA-36h1 for backup production. For such backup purpose, it is suggested to continue to maintain DKA cycles in the nearest future, but only for a reduced forecast lead time, (such as 6h), so that in case of repeated production interruption with

DKA37, DKA cycle can be extended to run full length equivalent to DKA37 for production delivery. Note that in order to apply such backup, it requires special arrangement with the production system at DMI.

3. **DMI-HIRLAM-SKA.** As long as DMI-HIRLAM suite is maintained with model results used for routine production, it is natural to define HIRLAM forecasts as the third substitute for DKA37 in case of delivery failure with HARMONIE system.

4. **ECMWF-forecast.** ECMWF global forecast of 4 cycle/day, with a native grid resolution of around 16 km, can be one of the candidate model substitutes.

Allocation of computation resource and typical execution time

Components	Involved job steps	nr of nodes in use	estimated execution time
LBC interpolation	ExtractBD, gl_bd	up to 50	150 to 200 s
Surface analysis	RunCarnari, OI-main	2, 1	150 to 215 s
3DVAR	Runscreening, Runminim, Runblend	20, 20, 1	170 to 310 s
Forecast, 58h	Forecast	124	4190-4300 s

In the above quoted estimation about execution time of the different processes, the variation in the forecast step is believed to be mainly associated with the steps about flow-dependent physics computations (on-off processes concerning radiation, hydrometeors, condensation etc.), and the variation in the assimilation steps are mainly associated with varying amount of observation data, data quality check and minimisation convergence toward required minimum.

Taking into account availability of number of HPC nodes and time slot, a special “Time Critical” mode has been designed, with which, in the case of delayed cycles, forecast length of past cycles is limited to a minimum 3h so as to ensure timely delivery of next cycle. For verification, the data from previous cycles with shifted forecast length is used to provide realistic quality measure.

4. Conclusion and near future outlook

4.1 Summary remarks

With implementation of DKA37, the HARMONIE system at DMI starts to feature full data assimilation with **3DVAR** and **3-hourly cycling** for a rather large domain with high resolution. One novel feature in the current implementation, compared to the reference HARMONIE 37h1.2, is the configuration of the cycling structure around asynoptic base time.

Thanks to a more frequent data assimilation, the DKA37 with 3DVAR has shown indication of improved spin-up for moisture properties. This hopefully address some of the concerns regarding consistency of forecast for short range, which has often been experienced with an intermittent data assimilation. Apart from this, the benefit of using 3-hourly 3DVAR instead of the simple upper air blending has so far been limited based on comparisons using conventional verification. This may partly be attributed to the fact that, at present, data assimilation in DKA37 has not made use of high spatial and temporal resolution observation, such as those of remote sensing from satellite, radar and GPS, especially those with humidity information. The conventional data that enters DKA37 are typically of a quite coarse resolution, and may hence be insufficient to provide critical information needed for forecasting convective-scale systems with HARMONIE. Nevertheless, it is believed that, by operationalising DKA37 with 3-hourly 3DVAR already at this stage, it provides an operational HARMONIE RUC infrastructure at DMI, which will be beneficial to promote an intensified research and implementation about use of remote sensing data in RUC system. Already now, the remote sensing data are entering the DKA37 system in a passive mode, which implies a real time technical monitoring of data quality without affecting yet assimilation results.

From feasibility investigation (Yang, 2012a) it has been demonstrated that, with an HARMONIE cycling with 3DVAR using asynoptic base time, it can deliver meteorologically comparable forecast skills compared to those with normal base hour, but with a significant speed-up of the product delivery. The configuration with asynoptic base time also improves greatly the load balance and resource utilisation with the operational HPC cluster at DMI, providing more opportunities for the pre-operational testing of future NWP suites. All in all, implementation of the asynoptic cycled DKA37 with 3-h 3DVAR on the DMI HPC has been successful, with little side effect on other existing operational activities.

It is anticipated that DKA37 will in the near future be followed by a mesoscale data assimilation with an increasing complexity, especially with inclusion of remote sensing data that is critical to enable a timely and frequent prediction of high impact weather.

4.2 Near future outlook

Implementation of DKA37 involved extensive preparation and parallel tests. Due to time constraints, the initial operationalisation of DKA37 still leaves several aspects to be improved or corrected. To list some of the potential area of further work in the near future,

- Use of hybrid optimisation with both MPI and Open-MP;
- Additional conventional observation from the road safety system and snow depth observations from Danish manual stations;
- Inclusion of remote sensing data (radar radial wind and reflectivity, ground-based GPS, GPS radio occultation, ASCAT wind), some of them being active;
- Variational bias correction (VarBC) applied to ATOVS (AMSU-A, MHS), Aircraft observation, ground-based GPS;
- Assimilation tuning, such as that about structure function and scaling;
- Porting of DKA37 for other model domains, such as GLA (nuuk), or enlarged domain on top of both DKA and GLA;
- Additional output and post-processing parameters.

Some of the above listed aspects about data assimilation and domain change may involve quality assurance evaluation, the rest are mainly of technical nature. It is envisaged that, following initial launch of DKA37, follow-up updates of DKA37 may be arranged during the first half of 2013, but most of such will be done in incremental fashion.

Acknowledgement Colleagues at Centre for Meteorological Model-System (CMM) in DMI, especially Bjarne Amstrup, Bent Hansen Sass, Henrik Feddersen, Jens Havskov Sørensen and Leif Laursen are acknowledged for discussions on various issues. Advices from system and data assimilation groups at the HIRLAM-B programme, especially Ulf Andrae, Ole Vignes, Trygve Aspelien, Mariken Homleid and Magnus Lindskog have benefited our progresses during the course of this work.

References

Andrae, U., 2012: Status of the HIRLAM/HARMONIE systems. HIRLAM Newsletter, nr. 59, 2012.

Brousseau, Pierre et al, 2012: Recent research on data assimilation at Meteo France, Presentation at the HIRLAM all staff meeting and ALADIN workshop, Marrakesh, Morocco, 2012.

Lindskog, L., M. Dahlbom, S. Thorsteinsson, P. Dahlgren, R. Randriamampianina, J. Bojarova, 2012, ATOVS processing and usage in the HARMONIE reference system.

Vignes, O., 2011, A simple strategy to take advantage of boundary (ECMWF) forecast quality in Harmonie 3D-Var cycling, available to HIRLAM community at: https://hirlam.org/trac/attachment/wiki/HarmonieSystemDocumentation/lsmixbc.ppt

Yang, X., 2012a: Configuration of HARMONIE cycling with asynoptic base time. HIRLAM Newsletter, nr. 59, 2012.

Yang, X., 2012b: Report on pre-release meteorological evaluation for 37h1 in HIRLAM-B. HIRLAM Newsletter, nr. 59, 2012.

Appendix A: Information related to DKA37

1. Source repository

The complete source for DKA37, with exception of climate data base and some supplementary utilities, is in the HIRLAM subversion repository tagged on Dec 20 2012:

https://svn.hirlam.org/tags/dmi-harmonie-37h1.2

which has since been updated on May 24 2013 to dmi-harmonie-37h1.2_oper

At DMI HPC platform, the copy of the repository is placed on

[hpcopr/hpcdev]: /oprhome/aldopr/harmonie_src/dmi-harmonie-37h1.2

2. Experiment configuration

The experiment configuration and control directory for DKA37 is located in
[hpcopr/hpcdev]: /oprhome/aldopr/hm_home/dka37
which usually contain, as minimum, following files and directories

- config-sh: basic settings about HPC platform, build option, repository, data path, job submission template.
- St: launch script.
- sms/config_exp.h: experiment configuration.
- control: scripts for handling of job queue reservation.

3. Grib data

- **Model level data for each 3-h cycle,** which contains complete model states in 3D, consists hourly forecast in **FA** format, as well as output from surface analysis, upper air analysis. In addition, model states for soil parameters in **lfi** format, are available for 0h and 3h. Due to the need for large storage space, model level data is not kept for more than 2 days, with exception of one cycle of each day, 23 UTC, for which the basic data set necessary for cycling is archived to the permanent storage system.

- **Thinned GRIB data from production cycle**, which is further passed into NINJO data-base and visualisation system for use by duty forecasters, as well as by forecast data-base and other downstream applications
  - /gdbopr/94/80/FC/SF/RL/1400_-18700_14336_-1364_022_022/yyyymmdd
  - hh/lll : 2D parameters, mainly those of surface or vertically integrated quantities, hourly
  - /gdbopr/94/80/FC/ML/RL/1400_-18700_14336_-1364_022_022/yyyymmddhh/lll: selected 3D upper air parameters, hourly

4. Runtime result directory HM_DATA

The runtime result directory of DKA37 is located in /oprdata/aldopr/hm_home/DKA37/. The main files and sub-directories are listed below.

YYYYMMDD_HH: runtime cycle directory. After completion of cycle, which includes input data preparation (the script family of “MakeCycleInput”), assimilation and forecast (“Date”), post-processing (“Postprocessing”), the directory is cleaned. Existence of the directory indicate either ongoing or incomplete cycle.

archive: sub-directory which contain results for each cycle.

- archive/yyyy/mm/dd/hh/: result data directory. kept for 36h
- archive/extract: data for model diagnostics, such as those of processed observation data in ODB data format; extracted model value of key surface and upper air parameters interpolated to surface observation stations; monitoring graphics and data for web presentation

DKA37:sub-directory with runtime log files with data created for each step of data assimilation cycle. After each sub-steps of the ”family” (MakeCycleInput, Date, PostProcessing), the directory is cleaned with logfiles collected per individual family and moved into the directory “archive”.

ECMWF: sub-directory with the lateral boundary data separated according to the original basetime of the ECMWF forecast.

lib: sub-directory with copy of the source code, scripts and basic constant and coefficients data which are actually used in the run.

observations: sub-directory with observation bufr data used in the run.

harmonie.check: runtime log recording check of cycle status of ongoing run
harmonie.log: runtime log containing timing and status (submit, active, complete, abort) of job steps as defined by the minii-SMS scheduler.
mSMS.log: runtime log of mini-SMS scheduler detailing job id of each submitted job steps.
mSMS.pid: temporary file containing process id of the submitted mini-SMS job scheduler. The presence of the file act as a lock to prohibit duplication of miniSMS job scheduler.
status: automatically generated log recording completion time of analysis, forecast and gribdata for each cycle.

5. Monitoring information

DKA37 automatically generates large amount of monitoring and verification information, most of these in simple graphic form and presented in web format. The graphic presentation of forecast is available via
https://hirlam.org/portal/dmi/dka37
The link contains entry to following monitoring information

- Forecast graphics: presentation of forecast charts and animations for MSLP, surface wind, temperature, cloud and precipitation
- Cycle status: information about cycle status of operational model

- Observation verification, inter-compared with other NWP products available to DMI. Currently DKA37 is compared to DKA (36h1.4), DMI-SKA and ECMWF
    - Surface: observation verification for surface parameters
    - Surface_map: error maps for different area or station lists
    - Surface_scat: scatter plot of surface parameters
    - Surface_skill: skill scores for wind, cloud cover and precipitation
    - Surface_sign: scores with error bar
    - Temp: verification against TEMP data in the vertical.
    - Temp_prof: vertically averaged profile of verification statistics
    - Temp_sign: scores with significance level
- surfass: time series of the domain averaged surface analysis increment
- costfun: evolution of cost function during data assimilation
- cpu: monitoring of per step CPU cost for forecast steps
- norms: domain averaged statistics for selected diagnostic quantities such as rain rate, cloud fraction
- amsuaarea/amsubarea: monitoring of amsu data per channel
- obstat: evolution of time series of available observation data per cycle
- obs_usage: observation data coverage map for SYNOP, AO (Aircraft observation), SHIP and TEMP data, and increment histogram
- satarea: same as above but for satellite data

Appendix B: HARMONIE DKA37 model output parameters.

DKA37 generates during forecast integration hourly output in FA format from start time point (0 h) to the final step (37h for production cycle and 03h for intermediate assimilation cycle). In addition, the model generate output for soil parameters in a so-called LFI format for selected time point at start of model run. The full content of the model data output is listed in the webpage of the HIRLAM-B programme,
https://hirlam.org/trac/wiki/HarmonieSystemDocumentation/Forecast/Outputlist/37h1

The original model grid-mesh in DKA37 is defined on rotated lambert projection. In order to enable easy handling of output data by the usual post-processing and graphic software at DMI, an extra step is made to convert selected forecast parameters into usual GRIB-1 data format and on a rotated lat-lon grid. Note that during the conversion, the procedure also chops off the extension zone which consists of 11 outermost grid points on the north and east border of the domain, hence the output data grid is 789x589x65 instead of the original 800x600x65.

For convenience, the forecast data stream are separated into two types, **ML** (**M**odel **L**evel) and **SF** **(S**ur**F**ace level), for 3D and 2D parameters, respectively. The data are placed in
/oprdata/gribdatabase/94/80/FC/**ML/**/RL/1400_-18700_14336_-1364_022_022/
for ML data and
/oprdata/gribdatabase/94/80/FC/**SF/**/RL/1400_-18700_14336_-1364_022_022/
for SF data, respectively.

Table 3 and Table 4 show the content of the parameters for these data streams, respectively. The ones with bold font denote new additions or modifications compared to the earlier 36h1-based DKA. Note that in CY37, the GRIB parameter, level number assigned to rain and snow have been changed in the HIRLAM-B programme in order to be more consistent with WMO GRIB standard. In this connection, the new cycle also introduced time range indicator to separate instantaneous from multi-time range quantities.

Table 3 DKA37 GRIB data output stream for diagnostic parameters (SF)

Description	parameter	level	type	time range indicator	unit
MSLP	1	0	103	0	Pa
surface pressure	1	0	105	0	Pa
geopotential	6	0	105	0	m2 s-2
Surface temperature	11	0	105	0	K
T2m	11	2	105	0	K
max T2m	15	2	105	2	K
min T2m	16	2	105	2	K
Visibility	20	0	105	0	m
U10m	33	10	105	0	m s-1
V10m	34	10	105	0	m s-1
Q2m	51	2	105	0	kg kg-1
RH2m	52	2	105	0	%
V. int. of cloud water	58	0	200	0	kg m-2
total precipitation	61	0	105	4	kg m-2
snow depth w. equiv.	65	0	105	0	kg m-2
boundary layer height	67	0	105	0	m
total cloud cover	71	0	105	0	1
fog	71	2	105	0	1
convective cloud	72	0	105	0	1
low cloud cover	73	0	105	0	1
medium cloud cover	74	0	105	0	1
high cloud cover	75	0	105	0	1
V. int. of cloud ice	76	0	200	0	kg m-2
surf. inst. solar radiation	116	0	105	0	W m-2

surf. inst. thermal rad.	117	0	105	4	J m-2
pseudo S.I. cl.top.tmp.	136	0	105	0	
pseudo S.I. w.v.b.tmp	137	0	105	0	
pseudo S.I. w.v.b.tmp+c	138	0	105	0	
pseudo S.I. cl.w.reflec.	139	0	105	0	
precipitation type	144	0	105	0	1
CAPE	160	0	105	0	J kg-1
gust U10m	162	10	105	2	m s-1
gust V10m	163	10	105	2	m s-1
accumulated rain	181	0	105	4	kg m-2
accumulated snow	184	0	105	4	kg m-2
cloud base	186	0	200	0	m
cloud top	187	0	200	0	m
accumulated graupel	201	0	105	4	kg m-2

Table 4 DKA37 GRIB data output stream for 3D model states (ML)

Description	parameter	type	level	Time range indicator	unit
temperature	11	109	1-65	0	K
u component	33	109	1-65	0	m s-1
v component	34	109	1-65	0	m s-1
Q	51	109	1-65	0	kg kg-1
cloud ice	58	109	1-65	0	kg m-2
cloud fraction	71	109	1-65	0	1
cloud water	76	109	1-65	0	kg m-2
TKE	200	109	1-65	0	m2 s-2

Previous reports

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