HARMONIE-AROME 4D-Var

Jan Barkmeijer[1], Magnus Lindskog[2], Nils Gustafsson[2], Jelena Bojarova[2]
Roohollah Azad[3], Isabel Monteiro[4], Pau Escribà[5], Eoin Whelan[6], Martin Ridal[2],
Jana Sánchez-Arriola[5] Ole Vignes[3] Roel Stappers[3] Roger Randriamampianina[3]

[1] Royal Netherlands Meteorological Institute, De Bilt, The Netherlands
[2] Swedish Meteorological and Hydrological Institute, Norrköping, Sweden
[3] Norwegian Meteorlogical Institute, Oslo, Norway
[4] Portuguese Institute for Sea and Atmosphere, Lisbon, Portugal
[5] Spanish Meteorological Agency, Barcelona and Santander, Spain
[6] The Irish Meteorological Service, Dublin, Ireland

Abstract

The status of CY43 HARMONIE-AROME 4D-Var is presented. Central to the HARMONIE-AROME 4D-Var implementation is the multi-incremental approach and the current setup comprises two outer loops at six and three times the main forecast grid resolution of 2.5 km. This setup is flexible though, as is the length of the observation window. Results will be given for 4D-Var in comparison to 3D-Var for three domains, centered around Scandinavia, the Iberian Peninsula and the Netherlands. The starting point here is that both 3D-Var and 4D-Var are using identical observations types. Generally speaking it can be concluded that 4D-Var forecasts are superior. The paper also gives an extensive list of opportunities from which particularly HARMONIE-AROME 4D-Var can benefit. For example, increasing the observation set, improving the treatment of observations during the observation window and new algorithmic avenues.

1 Introduction

1.1 General

The development of HARMONIE-AROME 4D-Var has been a long running data assimilation project of the HIRLAM consortium. After its first working version back in 2008 (CY38h1.2) the 4D-Var suite has matured over the years and has now reached a level that it can be considered for operational use.

At the core of the HARMONIE-AROME 4D-Var implementation is the use of the multi-incremental variational approach (Courtier *et al.*, 1994). Hereby the nonlinearities of the forecast model and the observation operators are better handled through a sequence of iterative re-linearizations (Trémolet, 2004). By solving a series (i=1,...,N) of minimizations, each minimization producing an analysis increment δx_i, the final increment δx is obtained through setting $\delta x = \sum_{i=1}^{N} \delta x_i$. For the i-th minimization or i-th loop, the corresponding cost function J_i reads as

$$\begin{aligned} J_i &= J_i^b + J_i^o \\ &= \frac{1}{2}((\delta \mathbf{x_i} - (\mathbf{x_b} - \mathbf{x_i})^{\mathrm{T}} \mathbf{B^{-1}} (\delta \mathbf{x_i} - (\mathbf{x_b} - \mathbf{x_i})) \qquad (1) \\ &+ \frac{1}{2} \sum_{k=0}^{K} (\mathbf{H_k^i M_k^i} \delta \mathbf{x^i} - \mathbf{d_k^i})^{\mathrm{T}} \mathbf{R_k^{-1}} (\mathbf{H_k^i M_k^i} \delta \mathbf{x^i} - \mathbf{d_k^i}) \end{aligned}$$

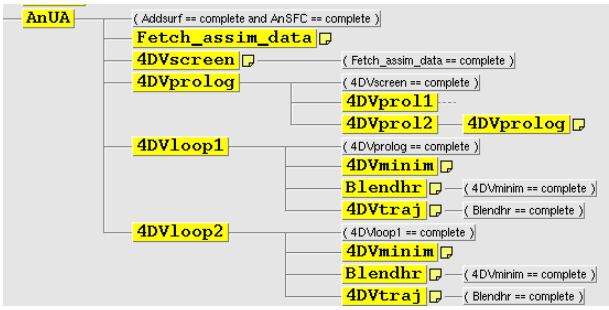

Figure 1: Illustration of the HARMONIE-AROME 4D-Var graphical workflow and the underlying script system, using two outer loops.

where x^b is the model background state valid at time t_0, $\mathbf{B}$ is the background error covariance matrix and the subscript k denotes time $t_k = t_0 + k\Delta t$ within the assimilation window, $d_k^i = y_k - \mathcal{H}(\mathcal{M}_k(x^i))$ are the innovations (observation minus background) where y_k denotes the observation vector at time t_k, $\mathcal{M}_k(.)$ denotes integration of the nonlinear model from time t_0 to time t_k and $\mathcal{H}(.)$ is the nonlinear observation operator, $\mathbf{M}_k^i$ and $\mathbf{H}_k^i$ denote the corresponding tangent linear (TL) model and TL observation operator respectively, linearized around the nonlinear trajectory $\mathcal{M}_k(x_i)$. In the first loop the model background state, x^b, is typically used as the first-guess $x^1 = x^b$ for the minimization. Eq. 1 is then minimized to yield δx^i, with $i = 1$. Following this, the first-guess state is updated with $x^{i+1} = x^i + \delta x^i$ ready to be used in the next loop.

In recent years HARMONIE-AROME 4D-Var has undergone several enhancements such as the possibility to assimilate an increased number of observation types (see Section 1.2), the use of spectral specific humidity and the use of a 3-hour observation window instead of 2-hour. The need for spectral specific humidity turned out to be essential in ensuring the tangent and adjoint models pass the so-called adjoint test ([TL$x, y$]=[x,ADy] for arbitrary model field perturbations x and y and with [.] the inner product used for deriving the adjoint model AD), a prerequisite for a well-defined minimization problem. Recently the 2-hour observation window version has been adapted to a 3-hour observations window. This in conjunction with the already 3-hourly cycling frequency ensures all observations from satellite overpasses are available for assimilation. The current HARMONIE-AROME 4D-Var further comprises two loops at 6 and 3 times the main forecast grid resolution of 2.5 km. This flexible setup allows for a minimization that is computationally affordable and with 10 and 15 iterations respectively most of the reduction in the gradient norm and cost function can be achieved (see Section 2.1). During the minimization use is made of the linear physics package developed for the ARPEGE and ALADIN models. It is comprised of linearized versions of some key physical processes. See Section 2.2 for a comparison between linear and nonlinear model evolution of increments in HARMONIE-AROME. Finally it is worth noting that a carefully designed scripting system together with the ecFlow graphical workflow package enables users to work with HARMONIE-AROME 4D-Var in a controlled manner, which is illustrated in Fig. 1 for a two outer loops set-up. At the end of each 4DVminim task the high resolution first-guess is modified by the computed low resolution analysis increment. After complementing with some extra model fields, using the Blendhr task, the trajectory run, 4DVtraj, propagates the analysis from the beginning of the observation window forward in time to produce the initial state for the high resolution forecast and updates the minimisation statistics in the ODB.

1.2 Observation usage

HARMONIE-AROME 4D-Var has been prepared to handle conventional (*in-situ* measurement) observation types (Table 1) as well as non-conventional types of observations (Table 2), including remote-sensing measurements. In the case of a 2-hour observation window for producing a forecast model initial state at hour HH observations from $HH - 60$ min to $HH + 59$ min are used. This will be changed to using observations from $HH - 60$ min to $HH + 60$ min, which is more appropriate. In the case of a 3-hour observation window for producing a forecast model initial state at hour HH observations from $HH - 120$ min to $HH + 59$ min

are used. Observations are sorted into seven and ten time-slots (for the 2-hour and 3-hour setup repsectively), sub-dividing the observation window in 10 min intervals at the beginning and the end and 20 min otherwise.

With respect to conventional observation usage z refers to geopotential height at mean sea level, $t2m$ temperature at the height of two metres, $rh2m$ relative humidity at the height of two metres, $u10$, $v10$ wind-components at 10 m, u, v wind-components, t temperature and q specific humidity. By default hourly SYNOP observations are used. There is also an option to apply a temporal thinning to use only the SYNOP observations valid at main analysis time. The reason for applying this temporal thinning is to avoid serially correlated observation errors that are presently not accounted for. These serially correlated observation errors can be caused by systematic observation errors at a particular station. In future HARMONIE-AROME 4D-Var versions a procedure for handling of these correlations (Järvinen *et al.*, 1999) will be considered.

Table 1: Usage of conventional types of *in-situ* measurements in 4D-Var.

Observation type	Variable
SYNOP	z
SHIP	z, u10, v10
DRIBU	z, u, v
AMDAR	u, v, t
PILOT	u, v
TEMP	u, v, t, q

For non-conventional observation usage as ztd refers to zenith total delay and std to slant total delay. RADAR observations are used with a temporal resolution of 1 h. The reason for not using RADAR observations with an even higher temporal resolution is again to avoid spurious effects of serially correlated observation errors not accounted for. Pre-processed RADAR observations of rh are derived by applying a 1-D Bayesian approach following strategies described by Caumont *et al.* (2010) and Wattrelot *et al.* (2014) using an observation handling in accordance with Ridal and Dahlbom (2017). It should be mentioned that SEVIRI radiances were assimilated in an earlier HARMONIE-AROME version (CY40) using hourly SEVIRI radiances in GRIB format. No problems are foreseen to introduce assimilation of these SEVIRI radiances also into the in the current (CY43) HARMONIE-AROME 4D-Var version. Observation types not yet tested in HARMONIE-AROME 4D-Var that are prepared for in HARMONIE-AROME 3D-Var are Global Navigation Satellite System (GNSS) Radio Occultations (RO) and Atmospheric Motion Vectors (AMV), radiances from the ATMS radiances and AMDAR humidity observations. No problems are foreseen with introduction of these remaining observation types into HARMONIE-AROME 4D-Var.

As in HARMONIE-AROME 3D-Var a Variational Bias Correction (VARBC) based on ideas of Dee (2005) is applied to handle systematic errors for satellite radiances and Ground-Based (G-B) GNSS ztds. Randriamampianina *et al.* (2011) studied VARBC configurations in the context of limited-area radiance assimilation. Following G-B GNSS ztd HARMONIE-AROME 3D-Var handling (Sánchez Arriola *et al.*, 2016; Lindskog *et al.*, 2017) one single VARBC predictor is applied by default. The use of multiple outer-loop iterations in HARMONIE-AROME 4D-Var imply a special handling of propagation of VARBC predictor coefficients in between different outer loops.

Table 2: Usage of of non-conventional types of observations in 4D-Var.

Observation type	Observed quantity
MODE-S	derived w and derived t
G-B GNSS	ztd, std
ASCAT	derived surface w
RADAR	derived rh, radial wind
AMSU-A, MHS, MWHS-2	microwave radiance
IASI, SEVIRI	infrared radiance

When it comes to specification of observation errors, identification of radiances contaminated by clouds, spatial thinning procedures and quality control there are no significant differences between HARMONIE-AROME 3D-Var and 4D-Var.

2 Demonstration of functionality

In this section HARMONIE-AROME 4D-Var functionality is briefly presented with a description of the cost function evolution, results from single observations tests and the consequences of enlarging the observation window from 2 hours to 3 hours.

2.1 Cost function evolution

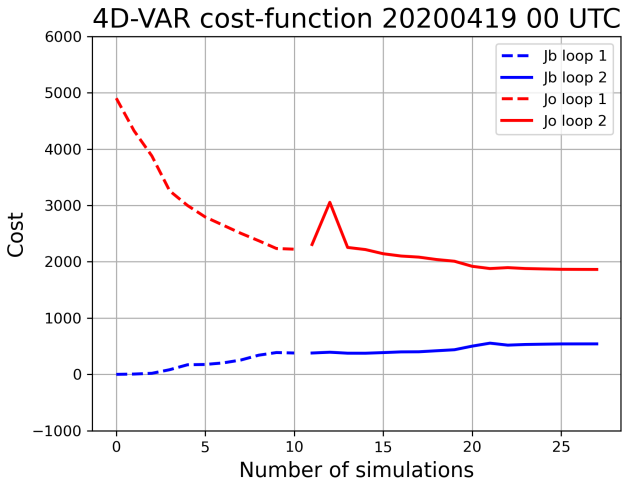

Figure 2: HARMONIE-AROME 4D-Var Jb and Jo parts of cost function for minimisation over MetCoOp domain for the 0000 UTC April 19, 2020 analysis using two loops at 6 and 3 times of the main forecast grid resolution of 2.5 km. For the first loop 10 iterations was used and for the second 15 (one iteration results in one or more simulations).

Figure 2 shows an example of the cost function, J, evolution during a HARMONIE-AROME 4D-Var minimization. The example is from the assimilation over a Nordic area using conventional as well as several non-conventional types of observations and a horizontal grid-distance of 2.5 km. Two outer loop iterations were used in the minimization of the cost function. Minimization in the first outer loop was applied at 15 km using a maximum of 10 iterations (15 simulations) and the minimization in the second outer loop was applied at 7.5 km using a a maximum of 15 iterations (20 simulations). Clearly the cost function values are mainly reduced in the first lower resolution loop, where larger scales in the background field are adjusted towards observations. The reduction of the cost function continues in the second, higher resolution loop, re-linearised around an updated nonlinear trajectory (see Eq. 1). It can be seen that at the end of the second outer loop the

cost function does not change much so that the minimization has reached a satisfactory level of convergence. For both loops a preconditioning of the minimization using the square-root of the background-error covariance matrix was applied (Courtier *et al.*, 1998; Gustafsson *et al.*, 2001). There is ongoing work aiming at applying a preconditioning by normalisation of the second outer loop using the Hessian of the cost function (Akkraoui *et al.*, 2012).

2.2 Single observations and linear vs nonlinear evolution

In order to further test the functionality of 4D-Var a single observation experiment was carried out for the 2-hour assimilation window 4D-Var. A specific humidity observation at 850 hPa from a radiosonde valid at the middle of the observation window was assimilated. Figure 3 shows the resulting increment at the beginning of the window at 850 hPa for this particular observation (no other observations were used in the data assimilation) and its linear evolution during the length of the observation window. The red dot in the middle panel indicates the location of the observation. In Fig. 4 the linear (top panel) and nonlinear evolution (bottom panel) of the increment in terms of specific humidity is presented at analysis time by means of a zonal cross-section through the observation location. The nonlinear evolution is given by the difference between two full physics HARMONIE-AROME runs starting from the first-guess and the first-guess modified by the increment. Clearly the propagated analysis produced by the trajectory run 4DVtraj (see Fig. 1) preserves the increment information to a large extent.

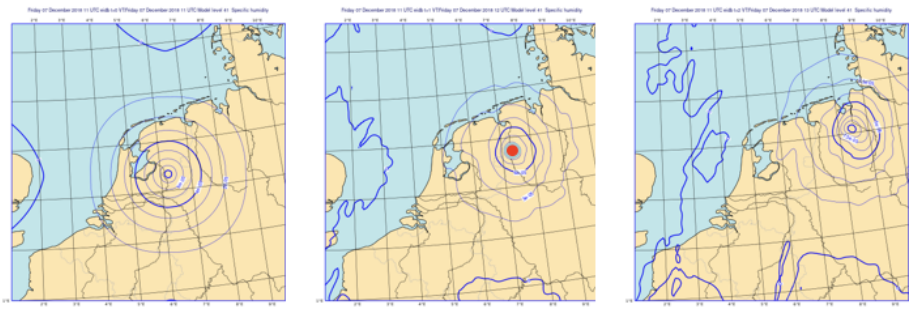

Figure 3: Linear evolution of the increment obtained from a 4D-Var single observations experiment. The 4D-Var observation window is 2 hours and the single observation, valid for the middle of the observation window, is for specific humidity at 850 hPa for a location denoted by a red dot. Panels show the increment at the beginning of the observation window (left), after 1 hour (middle) and after 2 hours (right).

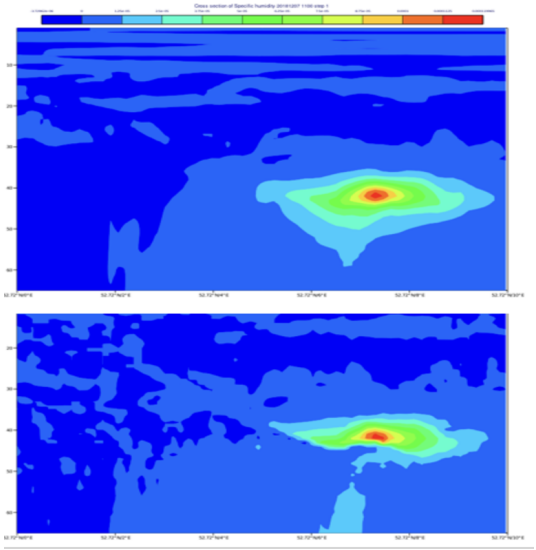

Figure 4: Cross section for specific humidity of the non-linear (top panel) and tangent linear evolution (bottom panel) of the increment shown in Fig. 3. The evolution is for 1 hour, which is the middle of the observation window (analysis time)

2.3 2-hour *vs*. 3-hour assimilation window

One of the recent enhancements of HARMONIE-AROME 4D-Var configuration was to apply a 3-hour assimilation window instead of 2-hour. The advantage of this configuration is shown for wind observations from ASCAT instruments on-board Metop satellites. Each day, these satellites overpass the Iberian domain between 09 UTC - 11 UTC and 21 UTC - 23 UTC. For the 2-hour window configuration, ASCAT winds between 10 UTC - 11 UTC and 22 UTC - 23 UTC are not considered. Fig 5 illustrates this point for 01 March 2020 at 12 UTC analysis. Although both configurations used observations from the two Metops available (Metop-B and Metop-C), for the 2-hour configuration, in contrast with the 3-hour one, only observations acquired between 11 UTC and 12 UTC were used in assimilation at 12 UTC analysis.

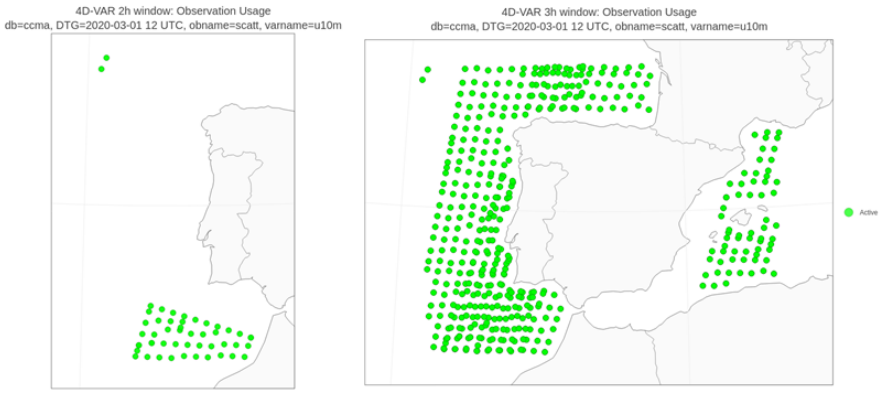

Figure 5: For 1 March 2020 ASCAT winds use in 4D-Var 2-hour window assimilation window (left) and 4D-Var 3-hour window (right) configurations, at analysis time 1200 UTC

3 Cooperation on three domains

3.1 Description of domains

An extensive evaluation of HARMONIE-AROME 4D-Var has been carried out for 3D-Var and 4D-Var by performing parallel data assimilation experiments over three different modelling domains centered around Iberian Peninsula (Iberia), the Netherlands (Netherlands) and Scandinavia (MetCoOp), respectively. The domains are illustrated in Fig 6. One advantage of running over a number of domains with different climates and characteristics is that this provides a more robust test of 4D-Var with respect to the typical local typical meteorological situations. It also prepares for a smoother local implementation in a pre-operational environment.

3.2 Experimental design

The evaluation of 4D-Var was partly through case studies investigating various aspects of HARMONIE-AROME 4D-Var, as demonstrated Section 2. We have also carried out a large amount of parallel experiments comparing 4D-Var with 3D-Var and also comparing various 4D-Var configurations over longer periods. Examples of such extended 4D-Var configuration comparisons are 2-hour *vs.* 3-hour observation window, impact of applying a Large Scale Mixing (LSM) of host model information into background state, utilisation of hourly SYNOP observations *vs.* utilisation at main analysis hour only, impact of applying the successive outer loop iterations at 15 km and 7.5 km *vs.* 15 km and 15 km, and assimilation of screen level $rh2m$ and $t2m$ observations. Here we will only briefly mention about some of these results but focus on three experiments comparing 4D-Var with 3D-Var. To make the comparison clearer LSM was not applied to either 4D-Var or 3D-Var in these experiments.

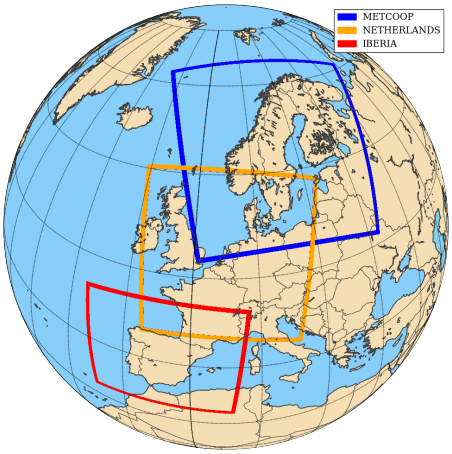

Figure 6: Domains used for the evaluation of HARMONIE-AROME 4D-Var.

The experiments were designed to use the same kind of observation in both the 4D-Var and the associated 3D-Var experiment. There were, however, some differences in configurations between the three different parallel experiments carried out. Each of the parallel 4D-Var *vs* 3D-Var experiments were carried out for a one-month period during 2020, covering three different seasons, and with 4D-Var configurations as summarized in Table 3. Before the start of each experiments a 10-day warm up-period was applied to spin-up surface fields and update VARBC-coefficients. The VARBC-coefficient files at the beginning of the warm-up period were taken from operational archive.

Table 3: Configuration of 4D-/3D-Var experiments carried out for 2020.

Conf	Iberia	Netherlands	MetCoOp
Period	07 Feb-07 Mar	14 Jun-14 Jul	20 Mar-20 Apr
Obs win (h)	3	3	2
Loop res (km)	15,7.5	15,7.5	15,7.5
Conv ob	all	all	all
Non-conv ob	ascat	mode-s ehs, ascat	amsu-a,mhs,iasi,ztd

4 Verification scores

Results from extended experiments carried out over the three domains reveal that 4D-Var produce forecasts that in general have better quality than the ones produced with 3D-Var.

For the experiment carried out over the MetCoOp domain 4D-Var produced clearly better atmospheric wind and humidity forecasts, as illustrated in Fig. 7. Temperature scores were rather neutral (not shown). The reason for the improved wind and humidity scores is most likely an improved use of wind and in particular satellite humidity-related observations taking the correct time of the observations into account and using the tangent-linear and adjoint of the forecast model within the minimization process. One problem with 4D-Var settings identified and solved while evaluating the coordinated experiments was that relaxation towards ECMWF host model information at the uppermost vertical levels was not applied in 4D-Var experiment, only in 3D-Var. This explains the worse 4D-Var forecasts at the uppermost levels seen in the left panel of Fig. 7.

Slightly worse scores for MetCoOp 4D-Var as compared to 3D-Var were obtained for short-range MSLP forecasts and standard deviation of forecasts of variables at two metre. This is illustrated in Fig. 8 for MSLP and rh2m forecasts. The reason for worse MSLP short-range forecasts is likely spurious wrap-around of surface

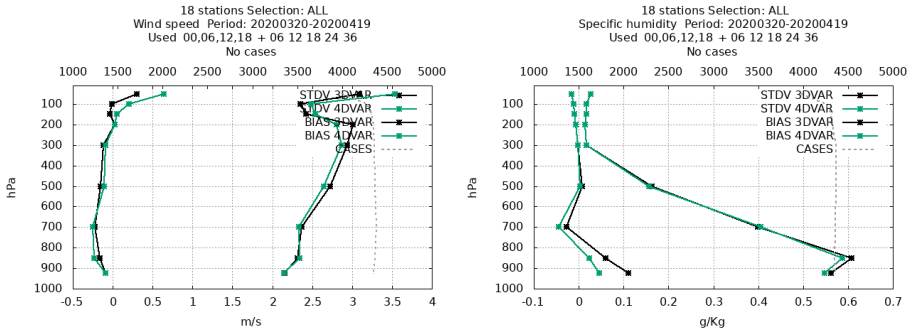

Figure 7: Verification statistics for MetCoOp domain, in terms of bias (BIAS) and standard deviation (STDV), of +06 to +36 h forecasts against radiosonde observations of wind speed (left, unit: m/s) and specific humidity (right, unit: g/kg) averaged over all observations within the domain and over the one month period. Scores are shown as function of vertical level and black full curves are for 3D-Var while green full curves represent 4D-Var. Grey dashed curve illustrates the number of observations used in the verification.

pressure analysis increments from one side of the domain to the opposite that can occur due to the use of bi-periodic spectral transforms in combination with large spatial surface pressure correlation length scales, a narrow extension zone with a relatively narrow redzone (in our case 150 km) and uneven distribution of surface pressure observations. The wrap-around problem is more pronounced in 4D-Var than in 3D-Var, because in 4D-Var the center of analysis increments may be projected significantly upstream of observation locations and well into the redzone area. This issue will be addressed in future versions by applying a wider extension zone. In addition there is a need for a revised blacklisting of z observations in MetCoOp 4D-Var as well as in 3D-Var. When it comes to relative humidity it can be seen that 4D-Var has slightly larger standard deviations than 3D-Var, in particular at short forecast ranges. However the bias is considerably smaller in 4D-Var. Regarding the standard deviations at two metre level, there is however room for improvement. With the Iberia experience in mind it is hoped that introduction of $rh2m$ and $t2m$ observations in 4D-Var as well as 3D-Var data assimilation will accomplish this.

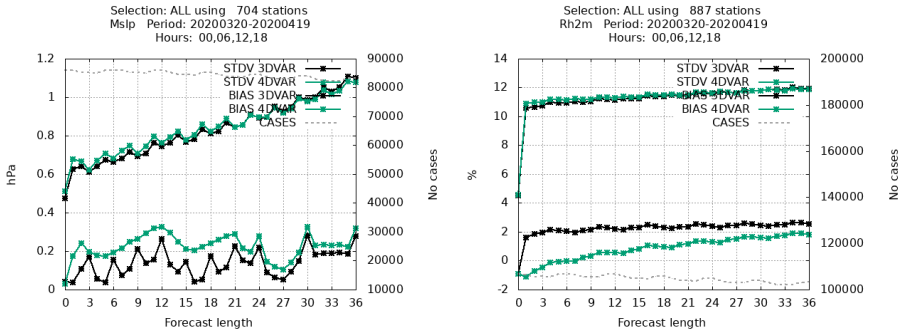

Figure 8: Verification statistics for METCOOP domain, in terms of bias (BIAS) and standard deviation (STDV) as function of forecast range for verification against SYNOP observations of $MSLP$ (left, unit: hPa) and $rh2m$ (right, unit: K) averaged over all observations within the domain and over the one month period. Black full curves are for 3D-Var while green full curves represent 4D-Var. Grey dashed curve illustrates the number of observations used in the verification.

Also for the experiments performed over Iberia, better wind forecasts were obtained with 4D-Var, as it is illustrated in terms of speed and direction in Fig. 9. In fact, comparing statistics for 3D-Var (black) and 4D-Var (green), improved scores are found for 4D-Var. In Iberia, apart from conventional observations, only ASCAT wind observations were used. A likely cause for the improved 4D-Var forecast scores is the 4D-Var ability to use these observations at the correct observation time.

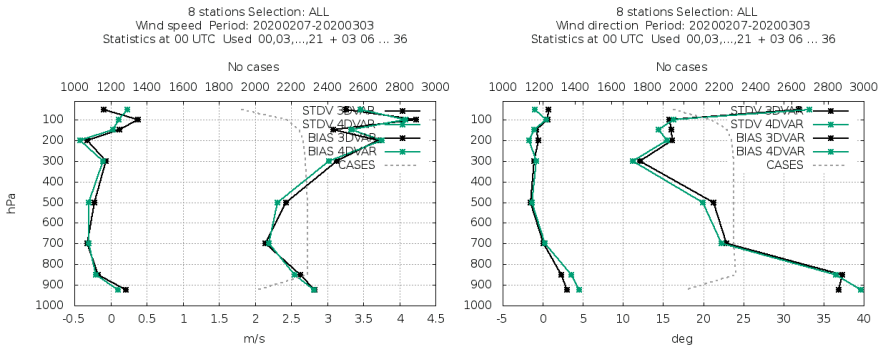

Figure 9: Verification statistics for the Iberian domain, in terms of bias (BIAS) and standard deviation (STDV) as function of forecast range for verification against radiosonde observations of wind speed (left, in m/s) and wind direction (right, in degrees) averaged over all observations within the domain From 07 Feb 2020 to 03 Mar 2020 period. Black full curves are for 3D-Var while green full curves represent 4D-Var (here the 3-hour configuration). Grey dashed curve illustrates the number of observations used in the verification.

For the Netherlands domain the performance for $MSLP$ improves with 4D-Var as compared to 3D-Var. This is most noticeable by a reduction of the bias, whereas the impact on standard deviation remains neutral, as is illustrated in Fig 10. The $MSLP$ forecasts do not seem to suffer from wrap-around issues as observed for the MetCoOp domain. With respect to screening level data the quality of $t2m$ forecasts improves as well, in particular for the bias. This is however not the case for $q2m$. Especially for the short range the 4D-Var forecasts are slightly worse both in terms of standard deviation as well as for bias, see Fig 10. Inclusion of $rh2m$ and $t2m$ observation in the future observation set may alleviate this.

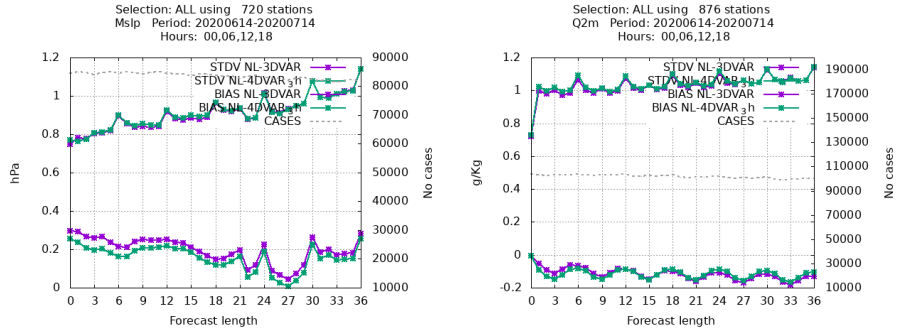

Figure 10: Verification statistics for the NETHERLANDS domain, in terms of bias (BIAS) and standard deviation (STDV) as function of forecast range for verification against SYNOP observations of $MSLP$ (left, unit: hPa) and $q2m$ (right, unit: g/kg) averaged over all observations within the domain and over the one month period. Purple full curves are for 3D-Var while green full curves represent 4D-Var. Grey dashed curve illustrates the number of observations used in the verification.

The large set of Mode-S EHS aircraft observations clearly has a positive impact on wind forecast performance in 4D-Var as shown in Fig 11. Note that in 3D-Var Mode-S EHS data is used only in a 30 minute interval around analysis time. Consequently roughly six times more Mode-S EHS observations are used in 4D-Var than in 3D-Var and at the appropriate observations time. In Fig. 11 scores for wind direction are shown indicating a decreased standard deviation and improved bias for 4D-Var. And although no additional humidity observations are assimilated other than from radiosondes, there is also a small improvement for 4D-Var specific humidity profiles as illustrated in Fig 11. This positive impact on humidity and associated parameters like cloud cover, fog and precipitation when using aircraft wind data has also been observed in earlier HARMONIE-AROME 4D-Var experiments and described in Gustafsson *et al.* (2018).

Additional sensitivity experiments carried out over the Iberian as well as over Netherlands domains revealed a benefit from using a using a 3-hour observation window as compared to using a 2-hour window. In addition, a positive impact from LSM, particularly for specific humidity was seen. This positive impact might decrease

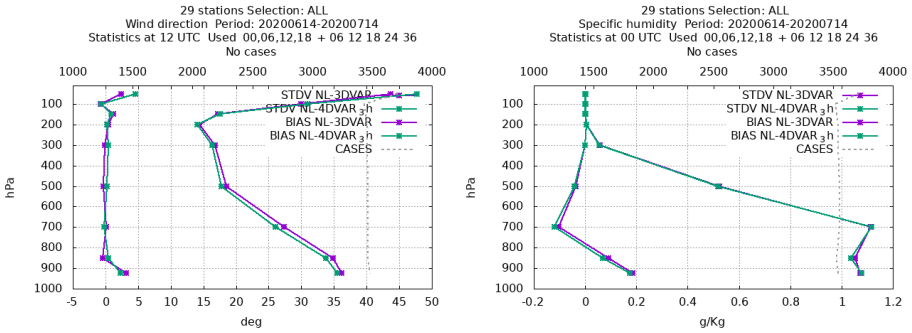

Figure 11: Verification statistics for the Netherlands domain, in terms of bias (BIAS) and standard deviation (STDV), of +06 to +36 h forecasts against radiosonde observations of wind direction (left, unit: m/s) and specific humidity (right, unit: g/kg) averaged over all observations within the domain and over the one month period. Scores are shown as function of vertical level and purple full curves are for 3D-Var while green full curves represent 4D-Var. Grey dashed curve illustrates the number of observations used in the verification.

when introducing assimilation of satellite-based radiances sensitive to atmospheric moisture, such as the MHS instrument. Finally a benefit of assimilating $rh2m$ and $t2m$ observations in HARMONIE-AROME 4D-Var experiments, in addition to their use in the surface analysis, has been demonstrated for Iberia.

5 Outlook

There are several ways in which HARMONIE-AROME 4D-Var can be further extended and improved. The various observation types which have not been explored yet provide an obvious route. We plan to extend the observation set with (not extensive) RADAR (both reflectivity and radial winds), AMDAR (humidity), GNSS STD, SEVIRI, AMV, MWHS-2, MHS, ATMS and on a longer term CRiS and MTG IRS.

Already mentioned are the introduction of a Hessian preconditioning to further speed up the minimization of the second outer loop, exploring the representation of serially correlated observation errors and adoption of a wider extension zone. Research versions with application of a wider extension zone (order of 50-100 grid-points) exists and will be further exploited. The larger extension zone framework requires more computer memory in minimization and also some transforms of domains (using the software tools GL or $FULLPOS$) in between 4D-Var and the surface analysis and forecast steps. It should be mentioned that the wider extension zone developments can be applied also in the context of 3D-Var. There have also been some early trials with weak constraint digital filter initialisation by introducing an additional Jc term to the cost function, penalising the divergent part of the flow-field. Longer term plans also include introduction of ECMWF simplified physics package and combining it with HARMONIE-AROME non-linear physics.

To benefit from high-quality large scale information provided by the ECMWF host model a spectral LSM might be applied to the background state (Müller *et al.*, 2017) in HARMONIE-AROME 3D-Var before upper-air data assimilation. Such a facility has also been introduced to 4D-Var, where the LSM is applied to the background state, valid at the beginning of the assimilation time-window and a positve impact have been seen over the Iberia. On the longer term one can as well think of introducing large scale host model information by introduction of additional constraints during the minimization of the cost function, as was done by Dahlgren and Gustafsson (2012). Another benefit from the large scale is the use of the optional relaxation to the ECMWF host model in the upper air levels (see 3D-Var experiment in Fig. 7.).

Based on ideas from HIRLAM 4D-Var (Gustafsson *et al.*, 2012) we are in the processes of enhancing HARMONIE-AROME 4D-VAR with an optional control of lateral boundary conditions. Some first results of controlling not

only the initial state but also the lateral boundary conditions (at the beginning of the observation window) are illustrated in in Fig. 12. Shown are surface pressure analysis increments for one particular cycle without and with control of lateral boundary conditions. One can see that introduction of control of lateral boundary conditions control effects the initial state analysis increments in particular close to the lateral boundaries, but also in the inner part of the domain.

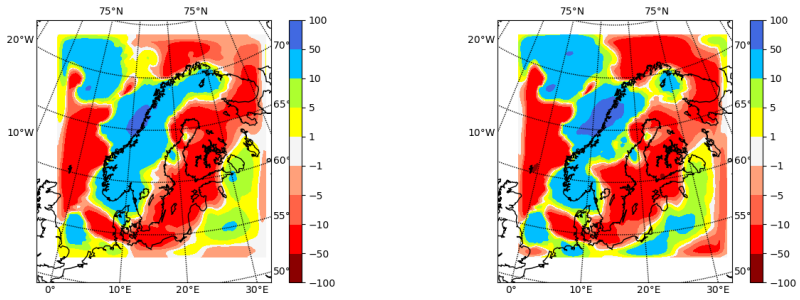

Figure 12: Two hour tangent-linear model development of surface pressure (unit: Pa) without (left) and with control of lateral boundary conditions. The propagation is for one particular assimilation experiment using one outer loop and 15 km assimilation grid distance and propagation of increments when convergence is reached.

One basic weakness of the default version of HARMONIE-AROME 4D-Var is the assumption of a perfect forecast model over the observation window – the forecast model is applied as a strong optimisation constraint. However, there exist possibilities to partly compensate for this weakness of 4D-Var. For example ECMWF has developed a method for controlling model errors during the data assimilation (Trémolet, 2006). This was in particular demonstrated beneficial in the stratosphere (Lindskog *et al.*, 2009), where it is now applied operationally at ECMWF. Related, the set-up of HARMONIE-AROME 4D-Var can also be used to determine a tendency increment that minimizes the cost function as given by Eq. 1 (Barkmeijer *et al.*, 2003). Then, instead of evaluating $\mathbf{M}_k^i \delta x^i$ we are interested in $\mathbf{M}_k^i \delta f^i$, where f^i is the tendency increment valid for loop i. In the current HARMONIE-AROME implementation the tendency increment does not change during the observation window, but is spatially varying both vertically and horizontally. In the case of opting for the tendency increment approach during the first outer loop, the subsequent nonlinear forecasts will apply a tendency increment for the appropriate lead time. For the forecast starting at analysis time this means that the tendency increment perturbation will not be active anymore after 1 hour into the forecast. Note that the part of the cost function associated with J_b remains constant during the minimization with respect to a tendency increment. In Fig. 13 the same single observation of specific humidity at 850 hPa valid for 12 UTC (middle of the observation window in this case) is used to determine the tendency increment. This tendency increment is plotted in the top left panel for specific humidity and differs for the analysis increment in the bottom left panel. It is slightly elongated and more located to the East. The middle panel slows the response of the linear model after applying this tendency increment for 1 hour. The maximum is located at the observation position and in fact J_o is zero. This linear response can be compared with the linearly evolved analysis increment as represented in the bottom row. The right panel shows the response at the end of the observation window.

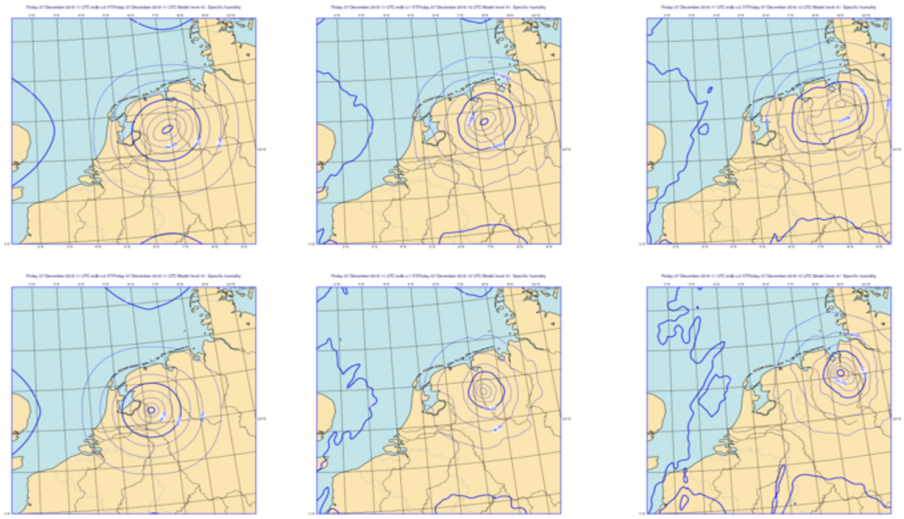

Figure 13: Top row. Tendency increment for the same single observation as used in Fig. 3 and its response after 1 hour (middle panel) and 2 hours (right panel) tangent linear integration. Bottom row. Identical to Fig. 3: the linear evolution of the analysis increment.

The main strengths of the 4D-Var scheme demonstrated so far comes from the possibility to assimilate observations at the proper observation time and from the flow-dependency of the analysis error structures due to tangent-linear and adjoint forecast model involved in the minimization process over a time window. The 4D-Var scheme will be further extended with the Hybrid EnVar functionality that allows for flow-dependent background error statistics. The Hybrid EnVar scheme based on an augmented control variable approach (Lorenc, 2003; Gustafsson *et al.*, 2014) has already been implemented in the HARMONIE-AROME forecasting system. It combines a full rank climatological background error covariance with ensemble-based flow-dependent evolution of the "error-of-the-day" uncertainty. The study to identify the best-choice initial conditions perturbations that are able adequately sample the "error-of-the-day" uncertainty for convective scale processes is on-going. On a longer time perspective we intend to apply a localisation including also hydrometeors (Destouches *et al.*, 2020) as developed for AROME 4D-Ens-Var and to enrich 4D-Var with a quasi-continuous overlapping window framework. We hope that nowcasting and very short range forecasting applications that are essential for modelling of convective scale processes will greatly benefit from this.

One particular challenge associated with initialisation of convective scale processes is different predictability time scales for different observed quantities. Temperature, wind and humidity profiles determine environmental conditions and have longer predictability time scales than for example cloud observations that bring in essential small scale details. An efficient data assimilation scheme on convective scales requires a long enough assimilation window to allow different model variables to influence each other and a short enough assimilation window to assure that controlled variables remain predictable over the window. The flexibility of HARMONIE-AROME 4D-Var and Hybrid EnVAR implementations allows to meet both requirements, which looks contradictory at first glance. An attractive feature of the Hybrid EnVar framework in the 4DVAR context is that the ensemble spans a space of non-linear model trajectories throughout that assimilation window. This provides flexibility to control some variables such as hydrometeors away from the start of the assimilation window.

Another idea is to develop for HARMONIE-AROME 4D-Var a tool to estimate the Forecast Sensitivity to Observation Impact (FSOI) with respect to short-range HARMONIE-AROME forecasts.

6 Conclusions

After several years of developments HARMONIE-AROME 4D-Var has now matured and is now at a stage where it can be considered for operational use. An added value as compared to 3D-Var has been demonstrated in terms of verification scores and there is ample opportunity for further improvement and applications of 4D-Var. In an operational framework it needs to be demonstrated that the benefit of 4D-Var remains when compared with a 3D-Var using LSM. In addition, computational aspects needs to be considered. The use of lower-resolution analysis increments, few iterations in the minimization of the cost function and only save the physics trajectory typically once every 30 minutes facilitate operational implementation. Some local optimization with regards to available computer capacity and observation cut-off will be needed.

7 References

Akkraoui, A.E., Trémolet, Y., Todling, R., 2012. Preconditioning of variational data assimilation and the use of a bi-conjugate gradient method. *Quart. J. Roy. Meteorol. Soc.*,**193**, 731–741. https://doi.org/10.1002/qj.1997

Barkmeijer, J., Iversen, T. and Palmer, T.N., 2003. Forcing singular vectors and other sensitive model structures. *Quart. J. Roy. Meteorol. Soc.*, **129**, 2401–2423, https://doi.org/10.1256/qj.02.126

Caumont, O., V. Ducrocq, E. Wattrelot, G. Jaubert, and S. Pradier-Vabre, 2010. 1d+3D-Var assimi- lation of radar reflectivity data: a proof of concept. *Tellus*, **62**, 173–187.

Courtier, P., Thépaut, J.-N. and Hollingsworth, A., 1994. A strategy for operational implementation of 4D-Var using an incremental approach. *Quart. J. Roy. Meteorol. Soc.*, **120**, 1367-1388.

Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F. and Fisher, M., 1998. The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part I: Formulation. *Quart. J. Roy. Meteorol. Soc.*, **124**, 1783-1808.

Dahlgren, Per and Gustafsson, N, 2012. Assimilating Host Model Information into a Limited Area Model , *Tellus A* , **64**, 15836 DOI: 10.3402/tellusa.v64i0.15836.

Dee, D., 2005. Bias and data assimilation. *Quart. J. Roy. Meteor. Soc.*, **131**, 3323–334.

Destouches, M., Montmerle, T., Michel, Y., Ménétrier, B., 2020. Estimating optimal localization for sampled background-error covariances of hydrometeor variables. *Quart. J. Roy. Meteorol. Soc.*, https://doi.org/10.1002/qj.3906.

Gustafsson, N., Berre, L., Hörnquist, S., Huang, X.-Y., Lindskog, M., Navascués, B., Mogensen, K. S. and Thorsteinsson, S. 2001. Three-dimensional variational data assimilation for a limited area model. Part I: General formulation and the background error constraint. *Tellus*, **53A**, 425–446.

Gustafsson, N., Huang, X.-Y., Yang, X., Mogensen, K., Lindskog, M., Vignes, O., Wilhelmsson, T. and Thorsteinsson, S., 2012. Four-dimensional variational data assimilation for a limited area model. *Tellus A*, **64**, 14985, DOI: 10.3402/tellusa.v64i0.14985.

Gustafsson, N., Bojarova, J. and O. Vignes, 2014. Hybrid Variational Data Assimilation for High Resolution Limited Area Model (HIRLAM). *Non-linear Processes in Geophysics*, **21**, 303–323, https://doi.org/10.5194/npg-21-303-2014.

Gustafsson, N., Janjić , T., Schraff, C., Leuenberger, D., Weissmann, M., Reich, H., Brousseau, P., Montmerle, T., Wattrelot, E., Bučánek, A., Mile, M., Hamdi, R., Lindskog, M., Barkmeijer, J., Dahlbom1, M., Macpherson, B., Ballard, S., Inverarity, G., Carley, J., Alexander, C., Dowell, D., Liu, S., Ikuta, Y. and Fujita, T. 2018. Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Quart. J. Roy. Meteorol. Soc.*, **144**, 1218-1256, DOI: 10.1002/gj.3179.

Järvinen, H., Andersson, E. and Bouttier, F. 1999. Variational assimilation of time sequences of surface observations with serially correlated errors. *Tellus*, **51A**, 469–488.

Lindskog, M. Dee, D., Trémolet, Y., Andersson, E. Radnóti and Fisher, M., 2009. Weak-constraint 4D-Var in the stratosphere. Quart. J. Roy. Meteorol. Soc., **135**, 695–706.

Lindskog, M., Ridal M., Thorsteinsson S. and T. Ning., 2017. Data assimilation of GNSS zenith total delays from a Nordic processing centre. *Atmos. Chem. Phys.*, 17, 13983–13998,
https://doi.org/10.5194/acp-17-13983-2017.

Lorenc, A. C., 2003. The potential of the ensemble Kalman filter for NWP—a comparison with 4D-Var. *Quart. J. Roy. Meteorol. Soc.*, **129**, 3183–3203.
https://doi.org/10.1256/qj.02.

Müller, M., Homleid, M., Ivarsson, K.-I., Koltzow, M., Lindskog, M., Midtbo, K.-H., Andrae, U., Aspelien, T., Berggren, L., Bjorge, D., Dahlgren, P., Kristiansen, J., Randriamampianina. R., Ridal, M., Vignes, O., 2017. AROME-MetCoOp: A Nordic Convective-Scale Operational Weather Prediction Model. *Weather and Forecasting*, **32**, 609—627. doi:10.1175/WAF-D-16-0099.1.

Randriamampianina R, Iversen T, and Storto A., 2011. Exploring the Assimilation of IASI Radiances in Forecasting Polar Lows. *Quart. J. Roy. Meteorol. Soc.*,**137**, 1700–1715. DOI:10.1002/qj.838.

Ridal, M. and Dahlbom, M., 2017. Assimilation of Multinational Radar Reflectivity Data in a Mesoscale Model: A Proof of Concept. *J. Appl. Meteorol. Clim.*, **56**, 1739–1751. https://doi.org/10.1175/jamcd-16-0247.1

Sánchez Arriola, J., Lindskog, M., Thorsteinsson, S. and Bojarova, J.. 2016. Variational Bias Correction of GNSS ZTD in the HARMONIE Modeling System. *J. of Applied Meteorology and Climatology*. **55**, 1259–1276.
doi:http://dx.doi.org/10.1175/JAMC-D-15-0137.1.

Trémolet, Y. 2004. Diagnostics of linear and incremental approximations in 4D-Var. *Quart. J. Roy. Meteorol. Soc.* **130**, 2233—2251.

Trémolet Y. 2006. Accounting for imperfect model in 4D-Var. *Quart. J. Roy. Meteorol. Soc.* **132**, 2483-–2504.

Wattrelot, E., O. Caumont, and Mahfouf, J.-F., 2014. Operational implementation of the 1D+3D-Var assimilation method of radar reflectivity data in the AROME model. *Mon. Weather Rev.*, **142**, 1852-–1873.