

Status and plans for the evaluation of Harmonie-AROME cy46 data assimilation tunable settings

J. Sanchez-Arriola, M. Lindskog, E. Whelan, J. Campins, B. Menetrier, D.Schönach, R. Eresmaa

46th EWGLAM & 31th SRNWP Meetings, 30th September-3rd October 2024 Prague

INTRODUCTION

The **HIRLAM-ALADIN** Research on Mesoscale Operational NWP in Europe-Application of Research to Operations at Mesoscale (HARMONIE-AROME; Bengtsson et al., 2017) limited-area Data Assimilation (DA) system used for operational NWP by several operational systems contain several tunable parameters influencing its performance. Recently, extensive DA experiments have been carried out over various domains and for different seasons to evaluate the performance of a new model version (cy46) as compared to the currently operational (cy43). Here we have analyzed the diagnostics of the data assimilation monitoring output from the new model version. For this analysis various objective methods described in the literature have been applied and utilizing existing available tools.

The latest cy46 version of this reference system, aimed for future operational use and described here was compared with cy43 version, on which currently used operational HIRLAM configurations are used. These parallel data assimilation and forecast experiments were carried out over three different domains corresponding to the currently operational domains of **MetCoOp** and **AEMET** and the planned operational domain of **UWC-west**. The geographical extent of these are shown in Figure 1.

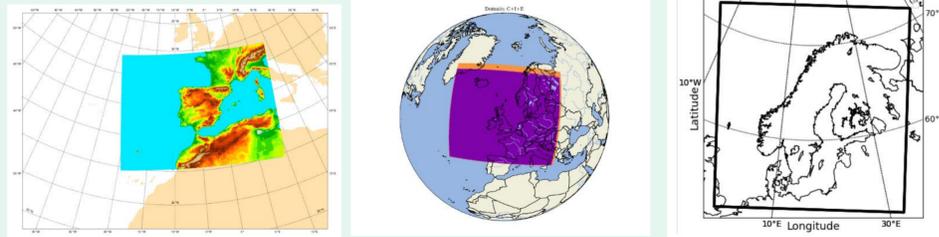


Figure 1 Model domains used in CY46 evaluation runs. AEMET (left), UWC-west (middle) and MetCoOp (right).

For each of these model domains one-month parallel cy43/cy46 deterministic data assimilation and forecast experiments were carried out for winter and summer seasons.

DA Diagnostic Tools

Tuning tools for diagnosing the functionality of the data assimilation from ODB feedback files have been prepared and documented to the ACCORD DA toolkit. (All available at <https://github.com/ACCORD-NWP/AccordDaTools>.)

2) Andersson & Järvinen. Andersson and Järvinen (1999)

A bad observation might deteriorate the data assimilation result, but excluding a good observation might result in detrimental quality.

Tool for estimating the Rejection Limits of the First Guess (FG) Check (example of **ob** in Figures 2 and 3):

$$((H(x_b))_i - y_i)^2 / \sigma_{b,i}^2 > FgLim \times \lambda \quad \lambda = 1 + \sigma_{o,i}^2 / \sigma_{b,i}^2$$

The input are the observation minus background (innovation) statistics from stored ODB files for various observation types and then produce histograms and transformed histograms. The transformation is done in such a way that when the transformed histogram starts to deviate from a linear variation then the innovation statistics starts to deviate from a Gaussian. The values where this happens are the limits that should be chosen.

The transformation is as follows:

$F = -2 \ln(f) / \max(f)$ where F denotes the transformed innovation histogram values and f denotes the original ones of rejections. $\max(f)$ denotes the number of rejections in the bin with maximum number of rejections.

1) Desroziers Desroziers et al. (2012)

The knowledge of errors of each piece of information used in the data assimilation is important for the performance

This tool is used for estimating background and observation error standard deviations (**ob** and **oo** respectively) to be used in the minimization of the 3D-Var penalty function. The input are the ODB feedback information regarding (1) observation minus FG departure, (2) observation minus analysis departure, (3) used **oo** values (4) used **ob**. The output is the information regarding vertical profiles of currently used and by Desroziers methods suggested **ob** and **oo** vertical profiles as derived from various observation types..

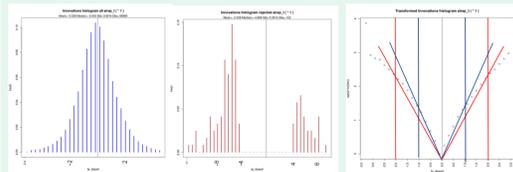


Figure 4. Andersson&Jarvinen. Left: Histogram of all ABO T obs. Middle:rejected ABO T by the FG Check, Right:Transformed histogram of all ABO T observations.

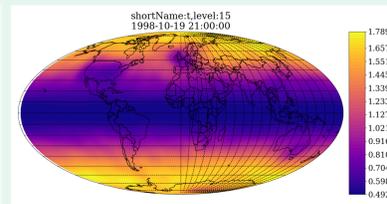


Figure 2. temperature background error standard deviations at model level15, roughly at 300 hPa vertical level (unit: K).

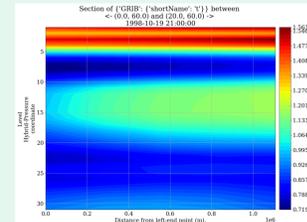


Figure 3. Vertical cross section of temperature background error standard deviations.

3) Ob stool. Bormann and Bauer (2010), Patrik Bucanek development.

In the lack of proper modelling of observation error correlations we handle by data thinning and inflation of oo. For thinning we need to approximate observation error correlation lengths for each instrument.

This tool is based on the Desroziers et al. (2012) method to estimate innovation correlations and separate these into background error correlations and observation error correlations using stored statistics of innovations and residuals at various distances from each other. This tool has been prepared to be used in this study but more work is needed to introduce all the satellite observations currently used by the three operational systems.

Background and observation error standard deviations

Shown here (Fig 5) as an example of results is upper air statistics from radiosondes to evaluate profiles of error standard deviations and for the different domains.

Vertical profiles of currently used and by **Desroziers diagnostics** proposed background error (**ob**) and observation error (**oo**) profiles for temperature, wind component and specific humidity are presented in Figure 5. The currently used **ob** profiles are derived using the tool Background Errors in Observation Space (BGOS), applied to each domain to get the background errors in observation space rather than control vector space. The **oo** profiles are taken from the ODB feedback files based on what is specified in the model source code.

- The most important difference noted between different configurations was that currently used upper-air variables **MetCoOp ob** have smaller values than **UWC-West** and **AEMET**, while **oo** values are similar between configurations. This implies that less weight is currently given in **MetCoOp** set-up than in **UWC-West** and **AEMET** to upper-air observations and surface geopotential observations, specially for humidity sensitive observations.
- In general, for existing error statistics we give more weight to **wind observations** than for **temperatures** from conventional types of observations within our assimilation system. And we see that the **oo** for radiosonde humidity varies with temperature, being high in summer season.

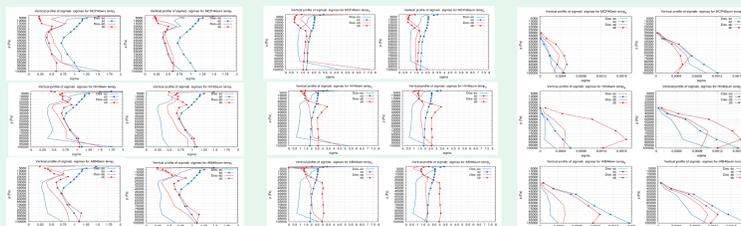


Figure 5: Vertical profiles of the currently applied full T (left) and wind (middle) and humidity (right) σ_b - (red dotted curve) and σ_o - (blue dotted curve) profile as well as by Desroziers method proposed σ_b - (red full curve) and σ_o - (blue full curve) profile. Results are for MetCoOp (upper), UWC-West (middle) and AEMET (lower) and for winter period (left) and summer (right). Results derived from ODB radiosonde feedback statistics.

Results

First-guess check rejection limits

To calculate the first guess rejection limits we applied the **Andersson & Järvinen method** to analyze each type of observation for each domain and period. The formula described above was applied to get the suggested First guess limits we should use to be more or less restrictive according with the histograms obtained for each case.

To summarize the results, two tables are shown with the currently used and the recommended values. Here shown as an example for upper-air conventional types of observations and GNSS and radar. Surface conventional observations and satellite base observations not shown here.

- For surface based observations we choose to focus on geopotential height from SYNOP, SHIP and DRIBU to start with. And conclude that the currently used rejection limits currently used by the three domains are more or less correct.
- For conventional types of upper-air observations we have addressed the limits for radiosondes (TEMP) and aircraft based observations (ABO) (Shown in Table1). Here, like for surface based observations, the currently used values are reasonably well confirmed by the diagnosed statistics.

- As suggested by Desroziers method there is signal that existing background errors for **MetCoOp** are of overall rather correct order of magnitude (REDNMC value), while **AEMET** and **UWC-West** ones should be slightly reduced.
- Desroziers method results generally imply that **oo** values should be overall reduced (and specially for radiosonde wind observations and SYNOP).

OBS/EXP	AEMET win/sum		MetCoOp win/sum		UWC-W win/sum	
TEMP T	5/5	4/4	5/5	4/4	5/5	4/5
TEMP u/v	4/4	3/4	4/4	3/5	4/4	2/3
TEMP q	4/4	2/2	4/4	2/2	4/4	5/3
ABO T	4/4	2/2	4/4	2/2	4/4	2/5
ABO u/v	4/4	3/3	4/4	3/3	4/4	3/3

OBS/EXP	AEMET win/sum		MetCoOp win/sum	
GNSS ZTD	4/4	2/2	4/4	2/2
RADAR RH	4/4	1/1	4/4	1/1
RADAR wind	1.6/1.6	-	1.6/1.6	1/1

Table1: measurements of Rejection Limit for the three domains. Left column: currently used, right column: recommended

- For GNSS ZTD and radar RH the currently used values are too high (see Table 1). That is confirmed by both **AEMET** and **MetCoOp** setups. It seems that for RADAR RH the reduction should be even larger than for GNSS ZTD. As a first step we propose to reduce GNSS ZTD and RADAR RH rejection limits from 4 to 2. In case of radar winds just **MetCoOp** used them for these runs and the recommended rejection limit could be 1 instead of 2
- For satellite based measurements the signal is a bit mixed and we choose here not to recommend changes in satellite rejection limit.

Conclusions

- An extended overhaul of the observation and background standard deviation used by the three operational domains for two seasons have been performed. Some differences in currently used values were found and highlighted like minor changes in REDNMC value used for some of the domains, or some reduction of the standard deviation errors values for some types of observations.
- The quality control rejection limits for First Guess Check have also been revised for the three domains and the two seasons and the conclusion has been that while for surface and upper air conventional observations, as well as for satellite based measurements, they seem to be adequate, for GNSS ZTD and RADAR RH a reduction of its values has been recommended.

Future work

- The last part of this extensive work is deriving thinning distances for all the operational domains and various types of observations, specially for satellite data. In addition we aim to improve and further develop the tools and evaluated impact of revised tuning parameters in model runs.
- Regarding the control of Rejection Limits for First Guess Check, in longer future we see room for improving the quality control procedure by using adoptive methods (VarQC, flow dep sigmab, AI/ML).
- The final conclusions and recommendations of this HARMONIE-AROME cy46 domain and season extensive comparison will be published in the coming months showing all the results.