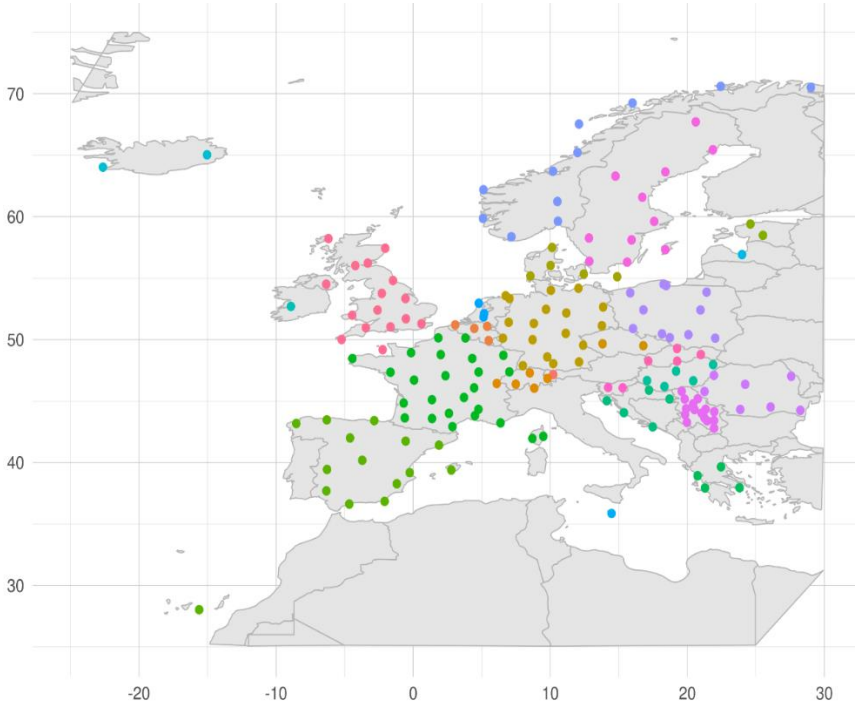
Regional Cooperation for
Limited Area Modeling in Central Europe

LACE
nwp central europe

Data assimilation and validation of radar radial winds observation

II



Author: Martin Petrovič (SHMU)
Host: Alena Trojáková (CHMI)
Purpose: Report from RC LACE stay
Place: CHMI, The Czech Republic, Prague
Date(s): 8.6.-21.6.2025 & 30.11.-13.12.2025

GeoSphere Austria · DHMZ · Czech Hydrometeorological Institute · IMGW · HungaroMet Hungarian Meteorological Service · METEO ROMANIA · SHMU · ARSO METEO Slovenia

Contents

List of Figures

Introduction

The assimilation of Doppler radar radial wind observations into high-resolution numerical weather prediction (NWP) models, such as ALADIN NWP system (Termonia et al, 2018) with its canonical model configurations ALARO and AROME, is a critical priority for enhancing the accuracy of short-range forecasts, particularly for convective events and local wind fields. Radial wind data provide direct measurements of velocity along the radar beam, offering high-frequency insights into atmospheric dynamics that other observation types cannot capture. However, integrating these data over Central and Southeastern Europe remains a challenge due to variability in data quality, complex error characteristics, and the necessity for robust preprocessing because of heterogeneous network of data providers.

This report summarizes the research and technical developments conducted during the RC LACE stay in 2025 at the Czech Hydrometeorological Institute (CHMI). Building upon the foundations laid during the 2024 stay, the 2025 internship included technical transition from ALADIN Cycle 46 to Cycle 48 and the deployment of the upgraded preprocessing tool, HOOF2.2py (version 1.10). A primary objective was to ensure dealiasing procedures and integrating more radars on domain.

Aim of this work involves the exploration of "degrees of freedom" within the data processing chain. We focused on spatial filtering methods (BATOR and Super-observations) and quality control. By utilizing the Block Bootstrap Anderson-Darling (BBAD) test, we quantified the Gaussian properties of the Observation-minus-Guess (OMG) departures to optimize the first-guess quality control rejection limits (ZFAC) used for quality control.

The results detailed in this report evaluate the trade-offs between data density and assimilation stability, providing a baseline for the future operational implementation of radar radial wind assimilation.

1. Work ground

The basis of the experiments were Doppler measurements of radial wind velocities obtained through the international OPERA (EUMETNET Operational Program on the Exchange of weather Radar information) network and the NIMBUS production line (EUMETNET,2024).

- **Domain:** Extended space including the national radar networks of the Czech Republic, Poland, Great Britain, Romania and Switzerland (Figure 1).
- **Data format**: The data was processed in HDF5 (ODIM) format, and we had to address structural differences in metadata (gain, offset) between individual providers (Czech Republic and Spain – thanks to Alena´s workaround, hack for missing attributes how/ and Nyquist velocity), especially when switching to the new NIMBUS line.

Radial wind provides unique information about atmospheric dynamics at high resolution, but its assimilation brings technical challenges:

- **Aliasing**: The phenomenon that occurs in the measured velocity exceeds the Nyquist velocity (NI), leading to incorrect interpretation of wind direction/velocity
- **Nyquist velocity**: We identified radars with critically low interval (NI < 30 m/s) that required increased attention in dealiasing.

Processing parameters (Degrees of Freedom)

In the experiments, we focused on the following configuration elements:

- **Dealiasing:** Algorithm for processing velocities using model wind projection.

- **Super-observations (SO):** Spatial averaging of data for better representativeness at the model scale.
- **Cleaner & Filter:** Physical cleaning of noise and isolated reflections in preprocessing.
- **ZFAC (Rejection Limit):** Adjustment of filter strictness during quality control (Background Check).

The connection of raw data with the assimilation system was ensured by the tool HOOF (Homogenization of OPERA Files) (Smerkol, 2020).

- The transition from version v1.9 (2024) to v2.2 (2025) was necessary for the stability of processing international datasets and the elimination of crashes caused by inconsistent HDF5 file headers.
- Data from HOOF were firstly homogenised and then converted to the ODB database by the BATOR system

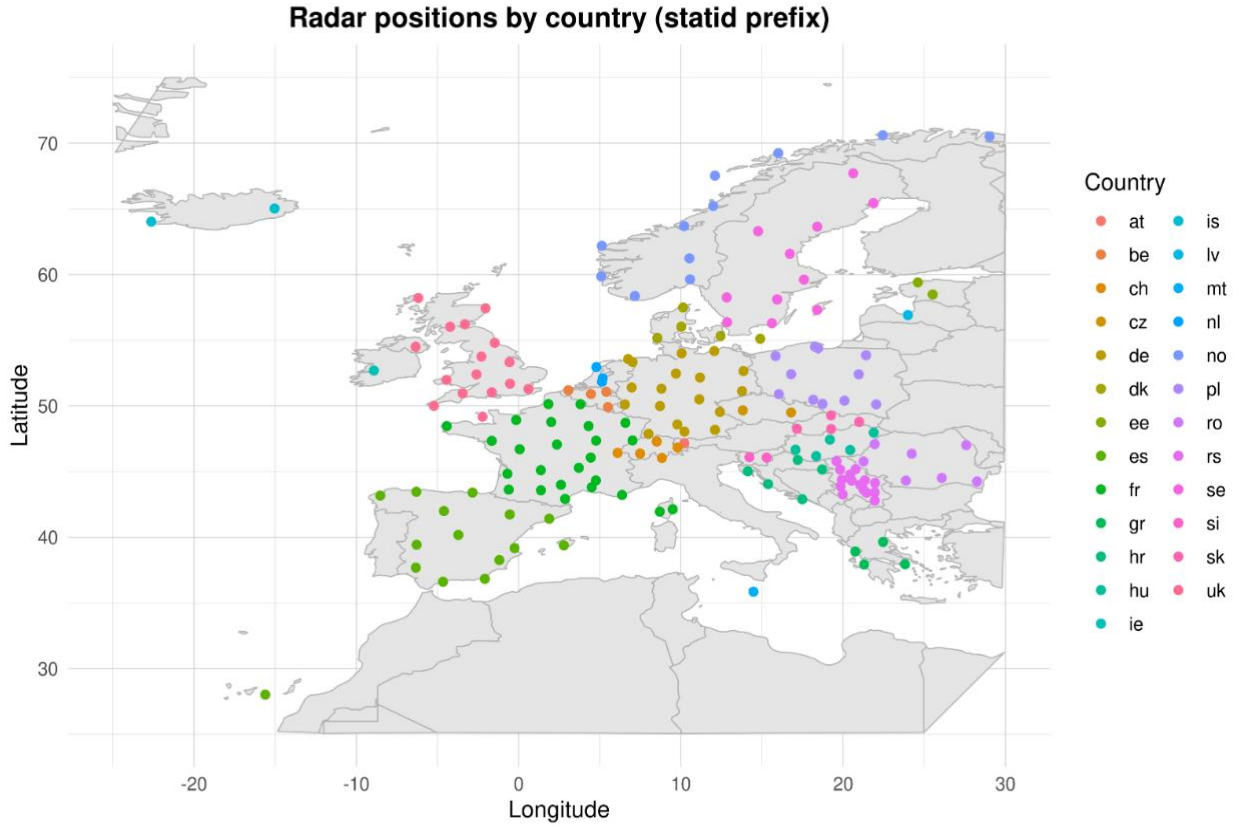


Figure 1 ***Position of European radars****, which are allowed to enter assimilation.*

2. Technical Setup

The main tool was the HOOF (Homogenization of OPERA Files) software, which provides format conversion and correction of physical measurement errors. The biggest challenge in assimilating radial velocities from Doppler radars is a phenomenon called aliasing. It occurs when the real wind speed exceeds the so-called Nyquist speed, causing the radar to measure the opposite direction or an unrealistic value. A more detailed explanation is in Sireci (2005), Brown and Wood (2007) and Petrovič (2024).

Methodology, according to Smerkol et al. (2025): We implemented an approach that uses the projection of the wind field from the model onto the radar beam. The model wind serves as an "initial estimate" that allows the HOOF algorithm to determine how many Nyquist intervals the measured velocity is shifted by.

HOOF Implementation
The transition to HOOF2.2.py was marked by several technical challenges. Subsequent migration to version 2.2 (v1.10) proved more stable, although specific attention was required for the handling of HDF5 file structures and the verification of dealiasing flags. When enabling superobbing in HOOF, it is necessary to change the *Superobbing output quality task name* or align it with what BATOR reads (*quality.superob* < = > *pl.imgw.quality.qi_total*).

BATOR Filtering vs. Super-Observations (SO)

A critical part of the technical setup was the evaluation of data thinning and filtering. We investigated the interaction between Super-Observations (SO) generated during the preprocessing stage (HOOF2.2.py) and the internal filters within BATOR.

I. Super-Observations principle

Superobbing is a procedure in which multiple nearby observation points are combined into a single representative point (super-observation; Ridal et al. 2024). The aim is to reduce the number of points entering the ALADIN model, thereby saving memory requirements and computing time while maintaining information value. The Superobbing "bin" is defined by factors, according to Smerkol (2020):

- **Range bin factor:** Number of cells in the radial direction.
- **Ray angle factor:** Number of rays in the azimuthal direction.
- **Arc size limit:** To prevent the grid from deforming too much with distance from the radar, the arc size in meters is limited. If the limit is exceeded, the edge rays are removed.

Radial velocity algorithm (VRAD):
For each bin, the system identifies all points with a measured value. If enough points are available (parameter *min percentage*) and their standard deviation is below the specified limit (parameter *max standard deviation*), the resulting super-observation value is calculated as the average of these points and the quality is set to 1. Otherwise, the bin is discarded (nodata).

II. Spatial filtering in BATOR

The BATOR system performs spatial filtering to eliminate noise and ensure smoothness of the data before assimilation. This process is performed in three steps through the subroutines *bator_radar_wind_cleaner* and *bator_filter_radar*. Quick resume from Montmerle and Faccani (2009) and report Petrovič (2024).

Bator_radar_wind_cleaner (Wind cleaner I – Elevation check)
This subroutine performs initial data cleaning at the elevation level using a median filter. It automatically rejects the entire layer if the proportion of valid observations falls below a critical threshold. By calculating the deviation of each central point from its neighbourhood's median, the system identifies and either adjusts or removes observations that exceed the defined error threshold.

Bator_filter_radar (Median filter)
It performs an in-depth analysis of each pixel by calculating the median from its defined neighbourhoods, effectively removing isolated "spikes" in the velocity field. This step is crucial for the numerical stability of the assimilation, as it filters out noisy data that does not meet the minimum number of points condition (COEFMEDIAN threshold).

Bator_radar_wind_cleaner (Wind cleaner II – Pixel check)
This step focuses on fine-grained cleaning at the pixel level, correcting or removing significant errors (threshold) by comparing them with the median of the surroundings.

3. Experimental design

To systematically evaluate the system, a series of passive experiments was designed.

A. Experiments on the Cy46 cycle

The first phase focused on isolating the impact of preprocessing steps in the HOOF tool and expanding the radar network:

- **t01 (Baseline):** Reference run with passive assimilation. Dealiasing and Superobbing (SO) were disabled.
- **t02 (Domain Expansion), cp -r t01/**: Expansion with international data from the network of Romania, Switzerland, UK and Poland.
- **t03 (Dealiasing), cp -r t02/**: Activating the dealiasing algorithm in HOOF on the extended domain.
- **t04 (Superobbing), cp -r t03/:** SO activation in HOOF.

B. Migration and Cy48 cycle experiments

- **t07 (Reference Cy48), cp -r t04/**: This experiment serves as a reference for the new cycle. BATOR has been recompiled with gmkpack to allow disabling filters directly via namelist.
  (*LBATOR_HDF5_CLEAN, LBATOR_HDF5_FILTER*)
  Settings: Dealiasing = ON, Superobbing = ON, Cleaner & Filter = ON.

- **t77 (Optimized setup), cp -r t07/:** Based on the results from Cy46, we disabled Superobbing in t77, but left dealiasing and filters active in BATOR.
  Settings: Dealiasing = ON, Superobbing = OFF, Cleaner & Filter = ON.

- **t08 (Without filters), cp -r t07/:** Identical to t07, but with filters turned off to verify the pure effect of geometric thinning.
  Settings: Dealiasing = ON, Superobbing = ON, Cleaner & Filter = OFF.

- **t09 (Modification of Rejection limit), cp -r t77/:** We have modified the fgchk.F90 source code to change the calculation of the ZFAC (rejection factor) parameter. The original fixed calculation was replaced by a dynamic relationship using the $RZFACT_{RADAR}$ variable, which can be controlled from the namelist:

  Origin t09/

$$ZFAC = \frac{400.0}{ZRIGHT * RBGQC(IVAR, IOBTYP, JFLAG)}$$

  Changed t09/

$$ZFAC = \left(\frac{400.0}{ZRIGHT * RBGQC(IVAR, IOBTYP, JFLAG)}\right) * RZFACT_{RADAR}$$

  Where *ZRIGHT* is scale factor (how much we trust model); *RBGQC* is Real Background Quality Control (which means standard deviation of background); *IVAR* is index of NVAR_DOW = 44; *IOBTYP* is index of observation type (for radars is 13); *JFLAG* is index of "actual state of observation" (active, passive, rejected).

4. Objective Quality Assessment

This chapter describes the methodology of radar data selection and the development of an automated system for evaluating the success of assimilation experiments. The main goal was to transform the raw model outputs into interactive statistical reports that replace the previous manual analysis using static histograms and inverse histograms.

The basic pillar is the initial evaluation of classical statistical characteristics such as the number of data (Figure 2), the distribution of departures between observations and guess (OMG, Figure 3) observation error (Figure 4), OMG histograms and also a visualization method that can be used to identify the exclusion limit (inverse histograms Anderson-Järvinen, 1999, Figure 5). To these characteristics we decided to add others such as skewness, kurtosis, data loss, coefficient of determination, a test for examining the normality of the distribution of the above-mentioned departures.

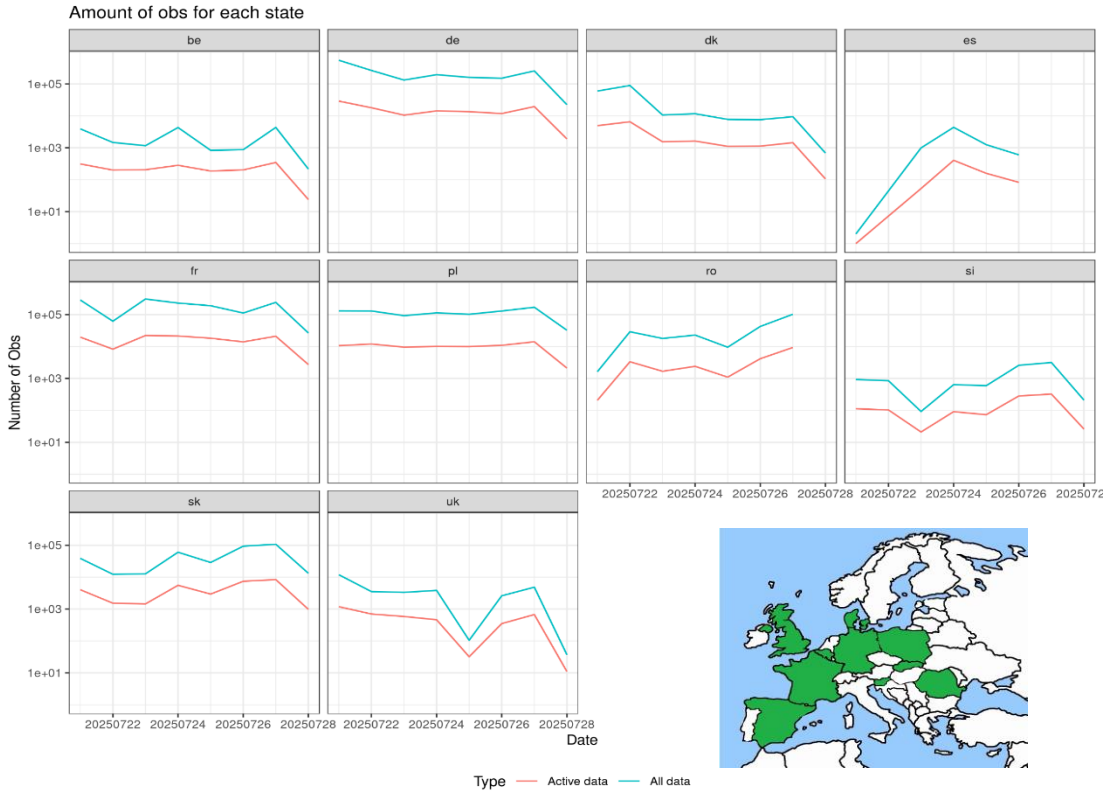


Figure 2 **Visualization of all and active Doppler wind observations from each European country** over a one-week period on log scale number of observations. The issue with unusual drops in observations counts (e.g. Romania - ro, Spain - es) were caused with missing TH for DBZ quantity.

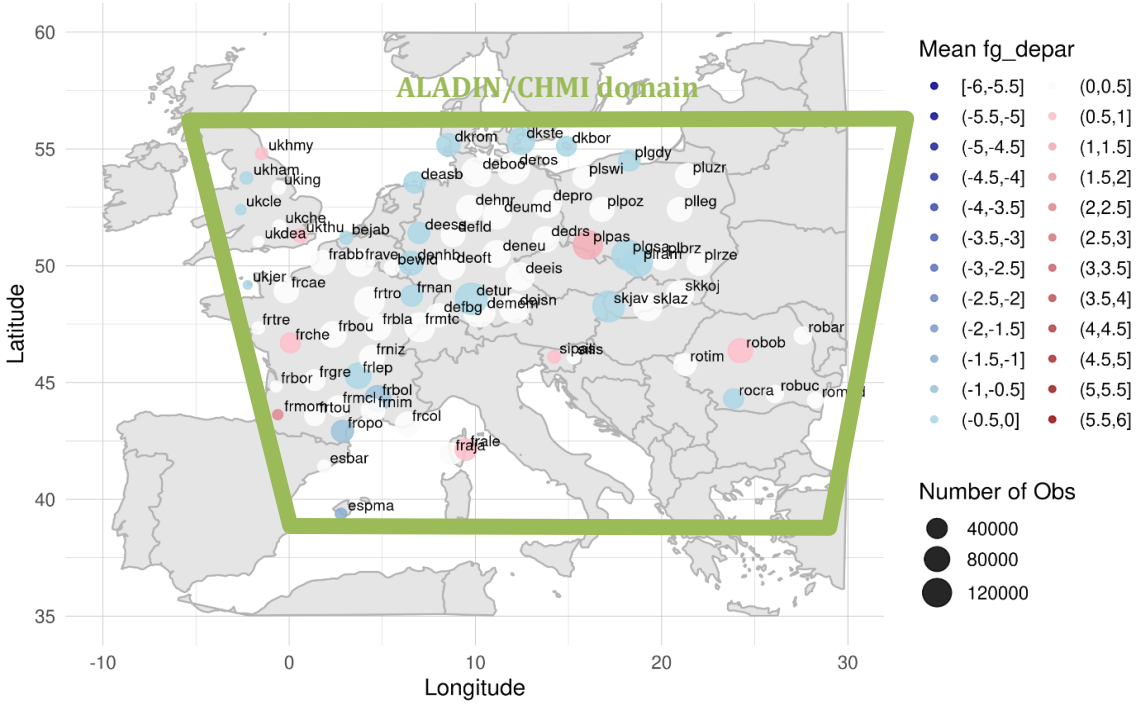


Figure 3 **Spatial distribution and monitoring first guess departures per radar from experiment t02.** The map displays radar locations within the ALADIN/CHMI domain (green). Radar stations are color-coded by their mean departure values, providing a spatial overview. The size of each point reflects the total observation count (Number of Obs), highlighting the density of data available for assimilation. Data from selected week, 21.-28.7.2025. The color coding of the stations does not indicate any significant regional biases.

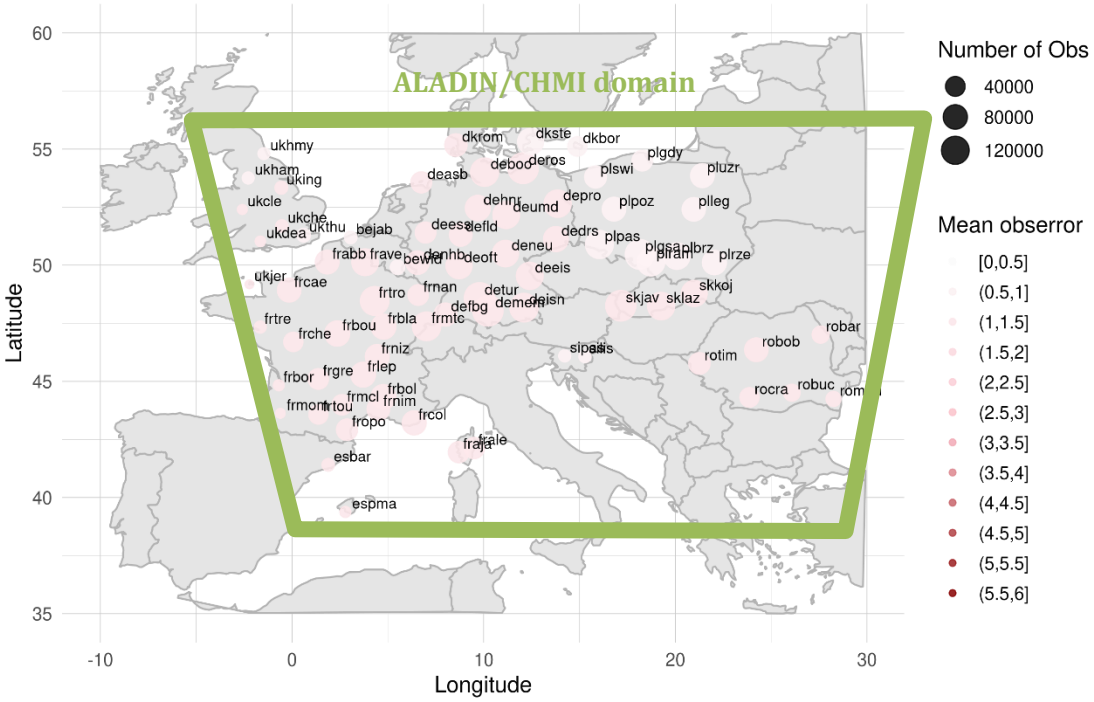


Figure 4 **Spatial distribution and monitoring mean observation error per radar from experiment t02.** The map displays radar locations within the ALADIN/CHMI domain (green). Radar stations are color-coded by their mean obs errors, providing a spatial overview. The size of each point reflects the total observation count (Number of Obs), highlighting the density of data available for assimilation. Data from selected week, 21.-28.7.2025.

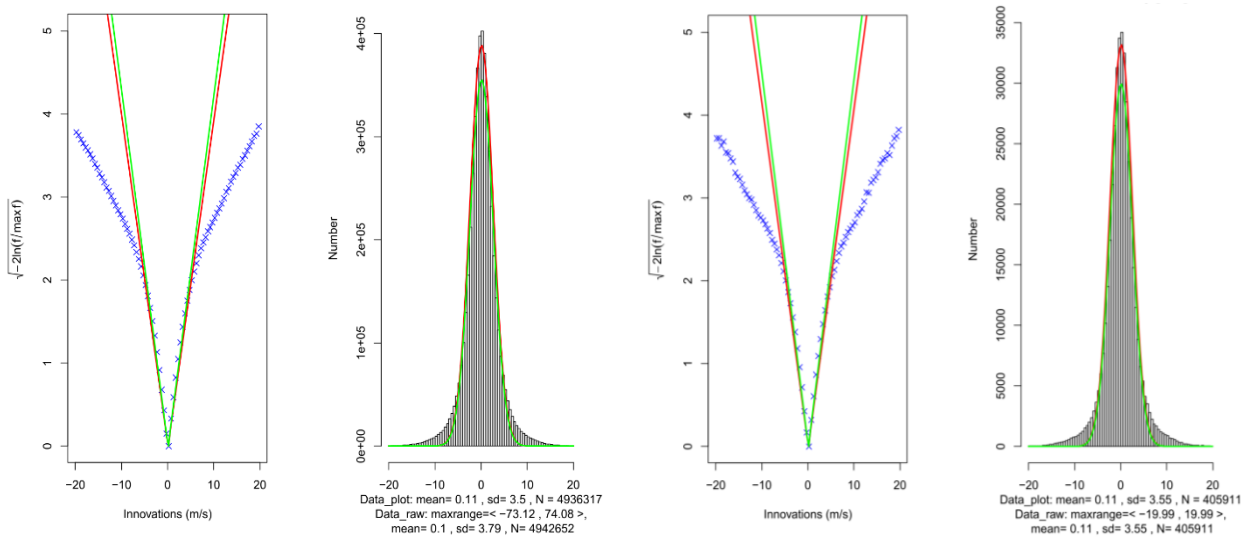


Figure 5 **Transformed histogram and histogram of OMG departures for experiment t02** of ALL data (left) and ACTIVE data (right). Red, or green, describes the normal distribution with offset, or without.

The left panels, on Figure 5, show the transformed histogram of observation departures for background, which have been transformed (Anderson and Järvinen, 1999; Petrovič, 2024). The red line describes the normal distribution (Gaussian curve) that corresponds to the mean and standard deviation of data, for better match was the line set offset to get better visual match. While green line is without offset. The blue crosses represent transformed bins from standard histogram, right panels.

However, a significant limitation of a generalized (domain-wide) evaluation is that it aggregates data points. To address this, it is essential to evaluate the statistical distribution at a higher level of granularity—at least per country and, ideally, per individual station ID (statid). This necessity motivated the usage of additional statistical metrics, such as skewness, kurtosis, data loss, and the coefficient of determination, which serve as a multi-layered cross check to verify the normality and quality of the departures across different spatial scales.

The Block Bootstrap Anderson-Darling (BBAD) test

In variational data assimilation, the innovation vector (OMG departures) is assumed to follow a Gaussian distribution. To objectively verify this assumption for the radar radial wind observations, we implemented a robust statistical framework based on the Anderson-Darling methodology.

The Anderson-Darling test (Stephens, 1974) is used to test if a sample of data came from a population with a specific distribution. It is a modification of the Kolmogorov-Smirnov (K-S) test and gives more weight to the tails than the K-S test does. This feature is critical in the context of radar observations, where non-Gaussian behaviour typically manifests as "heavy tails". Often caused by residual ground clutter, beam ducting, or non-linear observation errors that were not fully eliminated during the initial filtering stages. The K-S test is distribution free in the sense, that the critical values do not depend on the specific distribution being tested.

Standard normality tests, such as the basic Anderson-Darling (AD) test, are highly sensitive to sample size (N). With radar datasets containing hundreds of thousands of observations, even negligible deviations from normality result in a rejection of the null hypothesis. To address this, we utilized the Block Bootstrap Anderson-Darling (BBAD) approach (Vávra, 2015):

- **Resampling:** 1,000 data points were randomly sampled in contiguous 10-point blocks to preserve local spatial and temporal dependencies.
- **Iterations:** The process was repeated 100 times to calculate the median AD statistic.
- **Target:** An AD value below 1.50 was established as the threshold for "good" Gaussian fit.

The determination of the BBAD threshold values was based on a combination of theoretical properties of the Anderson-Darling normality test and back-calibration on real radar data. The theoretical limit for pure normality is unattainable for atmospheric measurements due to natural noise, and it was necessary to adapt the limits to reality. By analysing the histograms and inverse histograms of deviations (fg_depar) on the reference radars, we precisely identified the points at which the distribution still retains a Gaussian character. The limit of 1.50 (PASS) was thus determined as the point at which the histogram of pure data shows only negligible deformations. On the contrary, exceeding the value of 2.50 (FAIL) on the inverse histograms of radars was when bimodality and parasitic clutter begin to appear massively in the data.

Multicriterial optimization of the Rejection Limit

A critical task of the stay was to determine the optimal rejection limit for the screening process. Using the developed **Radar Quality Dashboards**, we performed a sensitivity analysis by varying the value of rejection limit. The tool analyses the sensitivity of innovations to changes rejection limit and searches for a "sweet spot" using the following parameters:

- **BBAD:** Block Bootstrap Anderson-Darling test
- **Skewness:** The primary goal is to achieve symmetry of the error distribution around zero. High skewness indicates the presence of asymmetric bias.
- **Kurtosis:** It focuses on minimizing “heavy tails” (outliers). The goal is to suppress extreme errors that deviate significantly from the Gaussian distribution.
- **R^2:** coefficient of determination of the linear fit between the deviation quantiles and the theoretical norm

Next two parameters are for informative purposes:

- **Data Loss %:** This is a critical constraint. An informative parameter to obtain "perfect" statistics at the cost of losing an unrealistic amount of data.
- **Sigma:** The development of error dispersion is monitored.

The overall status of the radar (OVERALL status) is determined by using a conservative approach: the resulting grade corresponds to the worst achieved metric (weakest link principle). This ensures that a radar with perfect skew, but extremely poor normality (BBAD test) will not be marked as satisfactory.

Metrics	*PASS*	*WARN*	*FAIL*	*Description*
BBAD	< 1.50	1.50 - 2.50	> 2.50	Normality test. Uses BBAD to normalize results for large datasets.
R^2	> 0.95	0.90 - 0.95	< 0.90	Goodness of fit for Gaussian distribution.
Skewness	< 0.50	0.50 - 1.00	> 1.00	Absolute value of distribution asymmetry.
Kurtosis	< 1.50	1.50 - 3.00	> 3.00	Absolute value of distribution sharpness

This automated scoring system allows you to immediately identify problematic radars across the entire OPERA network. If the radar is red, the tool provides clear information about whether the problem is kurtosis (too many outliers) or skewness (systematic bias), which significantly speeds up diagnostics and subsequent adjustments to the technical setup.

Analysis by Elevation and Distance

To move beyond global statistics, the statistical behaviour of radar observations was decomposed into elevation and distance bins. This spatial diversification is essential for identifying systematic errors, such as residual ground clutter at low elevations or beam broadening at long distances.

- **Global Performance:** The global analysis across the expanded domain (including new radar networks) confirmed that most of the integrated data meets the quality requirements for active assimilation (Figure 6).
- **Distance Effects:** A gradual decrease in R-Squared values was observed as distance from the radar increased beyond 120 km, reflecting the natural increase in sampling volume and beam broadening (Figure 7).
- **Elevation Effects:** Lower elevation angles (below 0.5°) showed higher Kurtosis values, likely due to residual ground clutter or beam ducting effects not fully removed by BATOR (Figure 8).

The diagnostic results can be visualized through interactive heatmaps within the dashboard. The heatmaps represent a matrix where rows correspond to spatial segments (e.g., elevation angles or distance categories like *NEAR* < 20 km, 20 km < *MID* < 60 km, 60 km < *FAR* < 120 km, 120 km < *EXTREME*) and columns represent the range of tested ZFAC limits. Each cell in the heatmap is color-coded (Green/Yellow/Red) based on the automated scoring system's "weakest link" evaluation for that specific bin and limit. The tool automatically marks the recommended rejection point with a sign of star, identifying the limit where the most spatial bins achieve a "PASS" status while minimizing data loss.

This granular approach allows us to pinpoint exactly why a radar might be failing. For instance, a radar might show excellent global statistics but reveal a "Red" status in the lowest elevation bin (0.5°) due to high Kurtosis, signalling that further clutter filtering or a stricter local RZFAC limit is required for that specific elevation.

App:
/home/mma274/app/radar_odb_check/rejection_limit_recommender/... Readme_RejLim_recom

Limit	BBAD	Kurt	Skew	R2	Sigma	Loss %
20	4.44	2.54	-0.05	0.977	2.92	0.0%
19	4.61	2.37	-0.02	0.978	2.92	0.0%
18	4.48	2.28	-0.01	0.979	2.91	0.0%
17	4.17	2.16	0.00	0.980	2.91	0.0%
16	4.05	1.98	0.01	0.981	2.90	0.1%
15	4.31	1.72	0.02	0.984	2.88	0.1%
14	3.79	1.58	0.01	0.985	2.86	0.2%
13	3.90	1.37	0.01	0.987	2.84	0.2%
12	3.10	1.19	0.00	0.989	2.82	0.3%
11	2.98	0.99	-0.01	0.991	2.79	0.4%
10	2.55	0.79	-0.01	0.993	2.75	0.7%
9	2.02	0.56	-0.01	0.995	2.70	1.0%
8	1.65	0.30	-0.02	0.996	2.61	1.7%
7	**0.95**	**0.05**	**-0.02**	**0.997**	**2.50**	**2.9%**
6	0.58	-0.20	-0.02	0.997	2.35	4.9%
5	0.78	-0.45	-0.03	0.994	2.13	8.5%
4	1.83	-0.68	-0.03	0.989	1.85	14.7%
3	3.85	-0.88	-0.03	0.980	1.51	24.8%

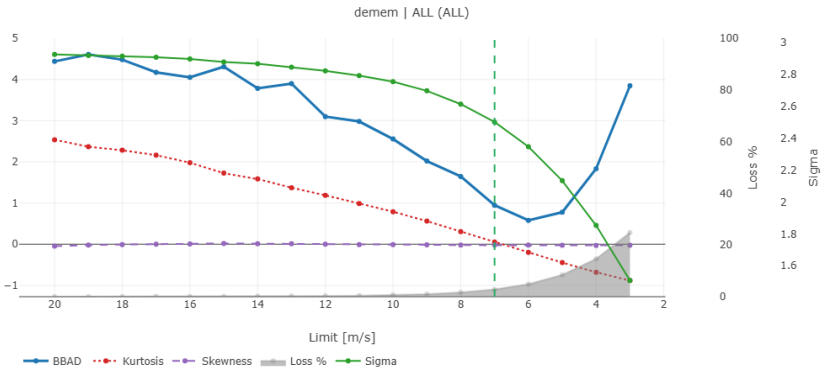
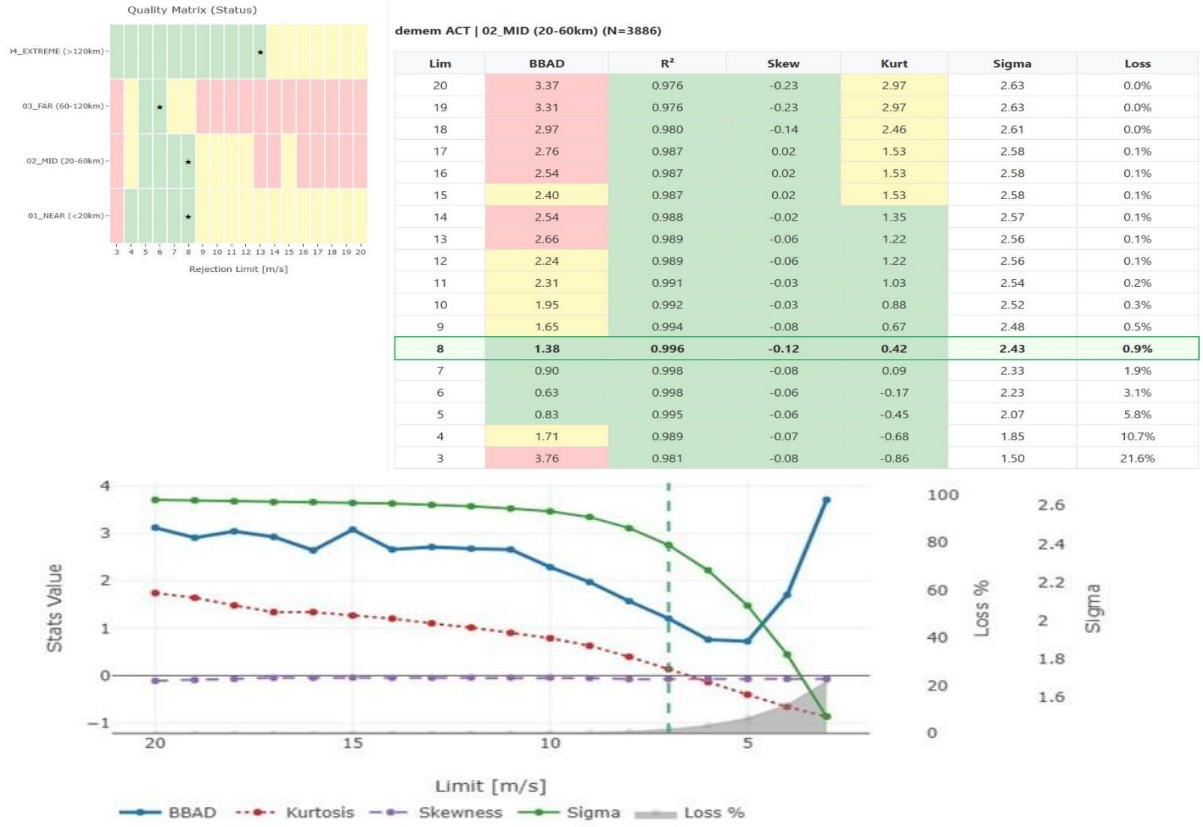


Figure 6 **Radar quality dashboard showing two panels of global** performance of DEMEN radar from experiment t02 (ACTIVE data).



Lim	BBAD	R^2	Skew	Kurt	Sigma	Loss
20	3.37	0.976	-0.23	2.97	2.63	0.0%
19	3.31	0.976	-0.23	2.97	2.63	0.0%
18	2.97	0.980	-0.14	2.46	2.61	0.0%
17	2.76	0.987	0.02	1.53	2.58	0.1%
16	2.54	0.987	0.02	1.53	2.58	0.1%
15	2.40	0.987	0.02	1.53	2.58	0.1%
14	2.54	0.988	-0.02	1.35	2.57	0.1%
13	2.66	0.989	-0.06	1.22	2.56	0.1%
12	2.24	0.989	-0.06	1.22	2.56	0.1%
11	2.31	0.991	-0.03	1.03	2.54	0.2%
10	1.95	0.992	-0.03	0.88	2.52	0.3%
9	1.65	0.994	-0.08	0.67	2.48	0.5%
8	**1.38**	**0.996**	**-0.12**	**0.42**	**2.43**	**0.9%**
7	0.90	0.998	-0.08	0.09	2.33	1.9%
6	0.63	0.998	-0.06	-0.17	2.23	3.1%
5	0.83	0.995	-0.06	-0.45	2.07	5.8%
4	1.71	0.989	-0.07	-0.68	1.85	10.7%
3	3.76	0.981	-0.08	-0.86	1.50	21.6%

Figure 7 **Radar quality dashboard showing panels of distance** performance of DEMEN radar from experiment t02 (ACTIVE data).

Lim	BBAD	R^2	Skew	Kurt	Sigma	Loss
20	5.27	0.977	-0.45	1.76	2.52	0.0%
19	5.19	0.977	-0.45	1.76	2.52	0.0%
18	5.03	0.977	-0.45	1.76	2.52	0.0%
17	4.75	0.977	-0.45	1.76	2.52	0.0%
16	4.94	0.977	-0.45	1.76	2.52	0.0%
15	4.90	0.977	-0.45	1.76	2.52	0.0%
14	4.90	0.977	-0.45	1.76	2.52	0.0%
13	4.63	0.980	-0.39	1.52	2.50	0.1%
12	4.63	0.980	-0.39	1.52	2.50	0.1%
11	4.25	0.982	-0.42	1.27	2.48	0.2%
10	4.12	0.984	-0.36	1.12	2.46	0.3%
9	3.92	0.988	-0.28	0.92	2.42	0.5%
8	2.08	0.993	-0.22	0.49	2.33	1.2%
7	1.56	0.996	-0.14	0.20	2.24	2.0%
6	**0.85**	**0.998**	**-0.08**	**-0.07**	**2.13**	**3.5%**
5	0.85	0.996	-0.02	-0.39	1.97	6.1%
4	1.44	0.991	0.04	-0.62	1.77	10.5%
3	3.29	0.983	0.02	-0.84	1.47	19.8%

Figure 8 **Radar quality dashboard showing panels of elevation** performance of DEMEN radar from experiment t02 (ACTIVE data).

5. Passive data assimilation experiments

Passive assimilation experiments represent a key step in linking the proposed statistical methods with the real operation of the ALADIN numerical model. In this phase, we detected innovations (OMG) from the OPERA network during a selected summer period in 2025, focusing on detecting technical barriers in BATOR, HOOF and Screening processes. The main goal was to verify the stability of the assimilation chain, examine the possible degrees of freedom that affect it, and obtain a statistically significant data sample for the final filter setting without the risk of affecting the quality of the forecast itself.

The period under review, from July 21, 2025, to July 28, 2025, included typical hot and dry conditions, and later in the week a cold front moved through, which became stationary over central Europe. Experiments were run with 3-hour assimilation cycle. Radar data are from Nimbus OPERA network, which were homogenised, dealiased and could have been filtered. European radars were used, specifically Belgium, Denmark, Germany, France, Spain, Slovenia, Slovakia, Poland, Romania and United Kingdom. Since several versions of the experiments were run and tested, the results from the latest setup version cy48t3mas_op1, which covered the ALADIN_CHMI domain and were run on the supercomputer of the Czech Hydrometeorological Institute (CHMI), are presented.

In the case of t07, when SO and Cleaner&Filter are applied, there is a significant negative data reduction. In this case, when HOOF SO and BATOR Cleaner&Filter are run simultaneously, they reduce the number of data from ~ 50,000 to 128, which represents almost 99.7% data loss. Such double thinning is a problem: SO aggregates nearby points into one super-observation, which changes the spatial density of observations. Subsequently, the BATOR median filter rejects the sparse field due to the small number of neighbouring observations (COEFMEDIAN threshold) and thus eliminates most of the observations.

The observation cost function contribution per observation (JO/n) is different for each experiment (and also per run), Table 1. In t07/, JO/n is between 4.4 -26.8, which is order of magnitude above normal. Only the most extreme outliers survived the double thinning, so the surviving observations per run have disproportionately large innovations. In case of t77/, JO/n ranges from 3.1- 12.3 and trends downward across run as the sample grows. While in exp t08/ swings between 4.1-13.

Exp t07/ reports ObsErr ~ 0.714–0.733, while t77/ and t08/ show ~ 1.09–1.14, Table 1. SO averages nearby raw observations into a single super-observation, artificially reducing its variance and thus its assigned representativeness error. The lower ObsErr in t07/ is therefore not a sign of better data it reflects the artificially smoothed character of super-observations, which do not represent point measurements and carry a different error structure. Native-resolution observations (t77/t08) correctly carry a higher representativeness error.

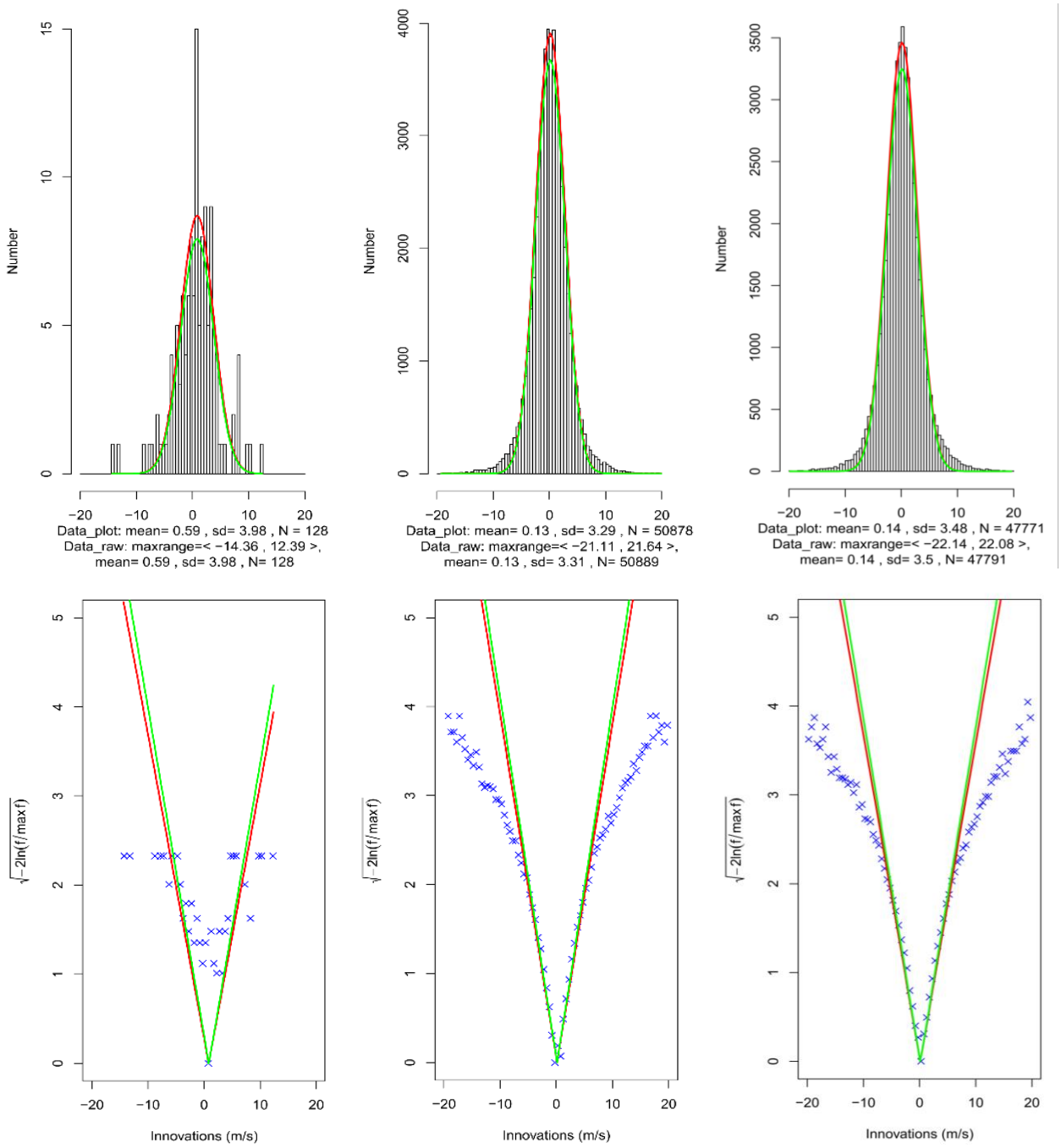


Figure 9 **Histograms (top) of OMG departures of active data for experiments t07(left), t77(middle) and t08 (right).** Transformed histograms for same experiments (bottom). Red, or green, describes the normal distribution with offset, or without.

Date	HH (UTC)	t07			t77			t08		
		DataCount DOW	JO/n	ObsErr	DataCount DOW	JO/n	ObsErr	DataCount DOW	JO/n	ObsErr
20250721	0	29	**15.92**	0.727	119090	**12.25**	1.13	31330	**13.02**	1.11
20250721	3	22	**7.89**	0.733	113734	**7.91**	1.14	33435	**8**	1.12
20250721	6	22	**24.54**	0.723	83859	**6.64**	1.12	28839	**7.75**	1.11
20250721	9	19	**4.4**	0.714	98558	**3.86**	1.11	34976	**4.93**	1.11
20250721	12	39	**10.26**	0.721	124500	**3.08**	1.11	42173	**4.11**	1.11
20250721	15	38	**10.89**	0.727	110417	**3.85**	1.12	38527	**5.02**	1.11
20250721	18	22	**19.02**	0.73	79243	**4.61**	1.11	27741	**7.82**	1.11
20250721	21	14	**26.83**	0.714	61788	**4.93**	1.09	22213	**7.44**	1.08

Table 1 ***Comparison of t07, t77, t08 experiments, focusing on data count, discrepancy with observations JO/n, and Observation error.*** *Noticeable differences were observed for one day assimilation. Experiments on cycle 48, where t07 (Dealias = ON, SO = ON, Cl & Fl = ON), t77 (Dealias = ON, SO = OFF, Cl & Fl = ON) and t08 (Dealias = ON, SO = ON, Cl & Fl = OFF).*

If we analyse runtime results, for one run is approximately for t07/, t08/ ~26min and t77/ ~16min. The fact that t77 is ~10 minutes faster despite processing 3× more observations than t08 is a strong operational argument. It directly contradicts the intuitive assumption that more data = longer run time. The bottleneck is HOOF2 preprocessing, not the assimilation itself, so eliminating SO gives you both slightly better quality and lower computational cost simultaneously - a rare win-win. Although is important to say that we don't use the newest optimized version of HOOF.

The aim of the series of experiment t09 was to quantify the impact of the $RZFACT_{RADAR}$ parameter on the quality and volume of assimilated radar data (DOW fg_depar) and to identify the optimal configuration. The testing was carried out in the mode with quality control filters (cleaner, filter) turned on and without the use of SO, which allowed to isolate the impact of the statistical rejection limit at the screening level.

At factor values of 1.0 and higher, the system exhibits identical statistical characteristics (N = 7,882, sigma = 2.48 m/s, innovation range -17.69 to +12.56 m/s). This confirms that above this limit the rejection limit ceases to be an active limitation and the fixed QC filters take over the main role. At low values (~0.01), there is a drastic reduction of data by more than 34% and an artificial trimming of departures to 2 m/s. Such a setting is unsuitable for operational practice, as it suppresses the real atmospheric signal and forces the model to rely too much on the first approximation (first guess).

$RZFACT_{RADAR}$	Amount N	Stand dev [m/s]	Mean [m/s]	Maxrange [m/s]
0.01	5 207	1.08	0.04	< -2.00, 2.00 >
0.1	7 730	2.20	0.12	< -6.30, 6.30 >
0.4	7 881	2.46	0.14	< -12.59, 12.56 >
1.0	7 882	2.48	0.14	< -17.69, 12.56 >
2.0	7 882	2.48	0.14	< -17.69, 12.56 >
10.0	7 882	2.48	0.14	< -17.69, 12.56 >

Conclusion

This work investigated the assimilation of Doppler radar radial wind observations (DOW) into the ALADIN cycle 48 NWP system during an RC LACE stay at CHMI in 2025.

The most significant finding of this work is the incompatibility between HOOF2 Super-observations and BATOR's internal spatial filters when applied simultaneously. In experiment t07, this double-thinning cascade reduced the active observation count from approximately 50,000 to fewer than 130 per assimilation cycle — a loss exceeding 99.7%. The mechanism is that SO reduces the spatial density of the input field below the minimum-neighbour threshold required by BATOR's median filter (COEFMEDIAN), causing entire elevation layers to be rejected. Disabling SO while retaining BATOR filtering (exp t77) resolves this conflict entirely, recovering ~100,000 observations per cycle with the tightest innovation distribution (sd = 3.29 m/s), the lowest bias (mean OMG = 0.13 m/s), and a 37% reduction in processing time compared to t07 and t08. The t77 configuration is therefore adopted as the recommended baseline for active assimilation experiments. However, further investigation in a fully active assimilation is required to determine whether the impact of disabling the cleaner/filter is only quantitative or whether it also produces qualitative changes on analysis.

The namelist controlled rejection limit ($RZFACT_{RADAR}$, experiment t09) provides a flexible mechanism for ZFAC tuning without recompilation. Based on the sensitivity analysis, $RZFACT_{RADAR} = 0.4$ is proposed as the initial value for next experiments. Below this threshold, meaningful atmospheric signal is suppressed; above 1.0, the parameter has no effect and fixed QC filters dominate. Full optimization of $RZFACT_{RADAR}$ per radar station using the BBAD dashboard is planned.

An objective quality assessment framework based on the Block Bootstrap Anderson-Darling (BBAD) test was developed and deployed. The framework evaluates radar data quality through a multicriterial scoring system (BBAD, skewness, kurtosis, R^2) applied across rejection limit sensitivity ranges, elevation bins, and distance bins. This replaces subjective visual inspection with an automated, station-level dashboard that directly informs the rejection limit configuration. The analysis confirms that residual non-Gaussianity in the innovation distribution is concentrated at low elevation angles (below 0.5°) and at ranges beyond 120 km, consistent with ground clutter and beam broadening respectively.

Several open points remain for future work. Active assimilation experiments are the necessary next step to assess the forecast impact of the t77 configuration. The representativeness error assigned to native-resolution observations requires tuning for the ALADIN_CHMI domain scale. Optimization of HOOF2 to a more recent version may reduce the per-radar SO overhead and partially close the runtime gap. Finally, evaluation of t09 results against the BBAD dashboard across the full European OPERA domain will provide a quantitative basis for setting $RZFAC_{RADAR}$.

Acknowledgement

I would like to thank Alena Trojáková for her professional mentorship and excellent working conditions at the Czech Meteorological Institute under the auspices of RC LACE. I also thank the team of the Numerical Weather Prediction Department of the Czech Meteorological Institute for their professional assistance and collegial approach.

References

ANDERSON, E., and JÄRVINEN, H., 1999. Variational quality control. *Quart. J. Roy. Meteor. Soc.*, 125, 697-722. https://doi.org/10.1002/qj.49712555416

BROWN, R. A., and WOOD, V. T. 2007. A guide for interpreting doppler velocity patterns: Northern hemisphere edition. *NOAA/National Severe Storms Laboratory*, 2007. Available: https://www.nssl.noaa.gov/publications/dopplerguide/Doppler%20Guide%202nd%20Ed.pdf.

EUMETNET, 2024. OPERA Available: https://www.eumetnet.eu/observations/weather-radar-network/.

MONTMERLE, T., and FACCANI, C. 2009. Mesoscale Assimilation of Radial Velocities from Doppler Radars in a Preoperational Framework. *Monthly Weather Review*, *137*(6), 1939-1953. https://doi.org/10.1175/2008MWR2725.1

PETROVIČ, M. 2024. Data assimilation and validation of radar radial winds observation. *RC LACE 2024.* From: RC LACE forum https://www.rclace.eu/media/files/Data_Assimilation/2024/repStay_MPetrovic_RadialWinds_CHMI_2024.pdf

RIDAL, M. et al. 2024. Optimal use of radar radial winds in the HARMONIE numerical weather prediction system. *Journal of Applied Meteorology and Climatology*, 2023, 62.12: 1745-1759. Available: https://journals.ametsoc.org/view/journals/apme/62/12/JAMC-D-23-0013.1.xml

SIRECI, O. 2005. Training course on weather radar systems. Module C : Processing basics in Doppler weather radars. WMO, *Turkish state meteorological service*. 2005. Available: https://www.yumpu.com/en/document/view/5300881/module-c-processing-basics-in-doppler-weather-radars-wmo .

SMERKOL, P. 2020. Documentation for the Homogenization of Opera files (HOOF) tool. *RC LACE 2020.* From: RC LACE forum https://www.rclace.eu/forum/viewtopic.php?p=2478&hilit=hoof#p2478 .

SMERKOL, P. et. al. 2025. Validation of torus mapping dealiasing method on Doppler velocities for use in numerical weather prediction. Online. *Meteorologische Zeitschrift*. 2025, vol. 34, no. 2, s. 109-123. ISSN 0941-2948. From: https://doi.org/10.1127/metz/1256.

Source code. Version cy48t3mas. Retrieved December 2025.

STEPHENS, M. A., 1974. EDF Statistics for Goodness of Fit and Some Comparisons, *Journal of the American Statistical Association*, 69, pp. 730-737.

VÁVRA, M., 2015. Testing for normality with applications. National Bank of Slovakia. ISSN 1337-5830.

Technical notes

LBATOR_HDF5_CLEAN, LBATOR_HDF5_FILTER

/home/mma274/local/pack/48t3mas_op1.01_BATOR_GNU/src/local/odb/pandor/module

- ***bator_module.F90***
  LOGICAL :: LBATOR_HDF5_CLEAN ! cleaner switcher
  LOGICAL :: LBATOR_HDF5_FILTER ! filter switcher

- ***bator_init_mod.F90***
  modified section USE BATOR_MODULE,
  !--------NAMELIST ----------

```
! initialisation namelist
!abuc
 LBATOR_HDF5_CLEAN  = .TRUE.
 LBATOR_HDF5_FILTER = .TRUE.
```

- ***bator_decodhdf5_mod.F90***
  modified section USE BATOR_MODULE, ONLY

```
LBATOR_HDF5_CLEAN=",LBATOR_HDF5_CLEAN,"
LBATOR_HDF5_FILTER=",LBATOR_HDF5_FILTER,"

! Spatial Filtering of DOW
!-------------------------
!    do l=1,NbSelectedElangles ! fileltering for each elevation
!      if (associated(Radar%FinalElev(l)%VRAD)) then
!        call
bator_radar_wind_cleaner('DOPW',ztent,ztwag,nbw,ilw,l,NbSelectedElangles,Radar%NRayons,Ra
dar%NPoints,.TRUE.,empty)
!        if (.not.empty) then
!          call
bator_filter_radar('DOPW',ztent,ztwag,nbw,ilw,l,NbSelectedElangles,Radar%NRayons,Radar%NP
oints)
!          call
bator_radar_wind_cleaner('DOPW',ztent,ztwag,nbw,ilw,l,NbSelectedElangles,Radar%NRayons,Ra
dar%NPoints,.FALSE.,empty)
!        else
!          write(batout,*) 'Wind data not avalaible. ELEVATION REJECTED'
!        endif
!      endif
!    enddo ! filtering all elevations

    do l=1,NbSelectedElangles ! filtering for each elevation
      if (associated(Radar%FinalElev(l)%VRAD)) then
      ! ---  CLEANER (KOLO 1) ---
        if (LBATOR_HDF5_CLEAN) then
          call
bator_radar_wind_cleaner('DOPW',ztent,ztwag,nbw,ilw,l,NbSelectedElangles,Radar%NRayons,Ra
dar%NPoints,.TRUE.,empty)
        else
```

```
          empty = .FALSE. !
        endif

        if (.not.empty) then
          ! --- FILTER ---
          if (LBATOR_HDF5_FILTER) then
            call
bator_filter_radar('DOPW',ztent,ztwag,nbw,ilw,l,NbSelectedElangles,Radar%NRayons,Radar%NP
oints)
          endif

          ! --- CLEANER ---
!          if (LBATOR_HDF5_CLEAN .and. LBATOR_HDF5_FILTER) then
          if (LBATOR_HDF5_CLEAN) then
            call
bator_radar_wind_cleaner('DOPW',ztent,ztwag,nbw,ilw,l,NbSelectedElangles,Radar%NRayons,Ra
dar%NPoints,.FALSE.,empty)
          endif
        else
          write(batout,*) 'Wind data not avalaible. ELEVATION REJECTED'
        endif
      endif
    enddo ! filtering all elevations
```

/home/mma274/local/pack/48t3mas_op1.01_BATOR_GNU/src/local/odb/pandor/namelist

- ***bator_namelist.nam.h***

NAMELIST / RADAR /
& LBATOR_HDF5_CLEAN, LBATOR_HDF5_FILTER

ZFAC / RZFACT_RADAR

/home/mma274/local/pack/48t3mas_op1.04_MASTER_REJLIM2/src/local/arpifs/module

- ***yomcosjo.F90***

```
! REJECTION LIMIT -----

REAL(KIND=JPRB) :: RZFACT_RADAR
```

/home/mma274/local/pack/48t3mas_op1.04_MASTER_REJLIM2/src/local/arpifs/namelist

- ***namcosjo.nam.h***

```
NAMELIST/NAMCOSJO/& RZFACT_RADAR
```

/home/mma274/local/pack/48t3mas_op1.04_MASTER_REJLIM2/src/local/arpifs/obs_preproc

- ***fgchk.F90***

```
USE YOMCOSJO    , & RZFACT_RADAR

!            3.2.27 DOPPLER WIND
  ELSEIF(ICMVNM == VARNO%DOPP) THEN
    IVAR=NVAR_DOW
```

```
!ZFAC = 400.0_JPRB/(ZRIGHT*RBGQC(IVAR,IOBTYP,JFLAG))  ! 20
m/s
   ! use adaptive rejection limit for Doppler winds
  ! ZFAC = 400.0_JPRB/(ZRIGHT*2.6_JPRB)
   WRITE(NULOUT,*) '>>>NEW RZFACT_USED:', RZFACT_RADAR
   ZFAC =
(400.0_JPRB/(ZRIGHT*RBGQC(IVAR,IOBTYP,JFLAG)))*RZFACT_RA
DAR
```

- ***defrun.F90***

```
USE YOMOBS   , & RZFACT_RADAR
RZFACT_RADAR = 0.4_JPRB
```