Regional Cooperation for
Limited Area Modeling in Central Europe

LACE
nwp central europe

Data assimilation and validation of radar radial winds observation

Author: Martin Petrovič (SHMU)
Host: Alena Trojáková (CHMI)
Purpose: Report from RC LACE stay
Place: CHMI, The Czech Republic, Prague
Date(s): 22.4. – 17.5.2024

Contents

List of Figures

Introduction

The study of radial winds using radar data has become increasingly vital for improving short-term weather forecasts, particularly within high-resolution numerical weather prediction (NWP) models like ALADIN. Radial wind data provide direct measurements of wind speed and direction along the radar beam, offering valuable insights into atmospheric dynamics. However, the integration of these data into operational models poses significant challenges due to the variability in data quality, coverage, and consistency across different regions.

EUMETNET Operational Program on the Exchange of Weather Radar Information (OPERA) produce among other products the quality-controlled volume data for NWP assimilation. An upgrade of its data centre (OIFS) to the new production line (NIMBUS) is planned for 2024 (EUMETNET, 2024).

We focused on addressing these challenges by comparing and validating two key radar datasets: OIFS and the newly introduced NIMBUS. The primary aim was to assess the potential of these datasets for improving the accuracy of weather forecasts through enhanced data assimilation techniques. This involved the application of spatial filtering methods to refine the data quality and the exploration of passive assimilation strategies to incorporate these refined datasets into the ALADIN model.

1. Basic overview of OPERA datasets: OIFS & NIMBUS

Our initial focus was to identify the differences between the provided OIFS and NIMBUS datasets. We aimed to compare the content of the data in the respective datasets from one randomly chossed term (01.04.2024). Firstly, we identified missing radar stations which were missing in OIFS (FRAJA, ISBJO, ISSKA, LVRIX) or NIMBUS (IEDUB, PLPAS, RSFRG) datasets. The Table 1 shows differences in the number of radars recorded in OIFS and NIMBUS databases for each country as of April 1, 2024.

Amount of radars (01.04.2024)					
Country	**OIFS**	**NIMBUS**	**Country**	**OIFS**	**NIMBUS**
Belgium (be)	3	3	**Malta (mt)**	1	1
Czech (cz)	2	2	**Netherlands (nl)**	2	2
Croatia (hr)	5	5	**Norway (no)**	10	10
Denmark (dk)	4	4	**Poland (pl)**	7	6
Estonia (ee)	2	2	**Portugal (pt)**	4	4
Finland (fi)	11	11	**Serbia (rs)**	1	0
France (fr)	21	22	**Slovakia (sk)**	4	4
Germany (de)	17	17	**Slovenia (si)**	2	2
Hungary (hu)	5	5	**Spain (es)**	13	13
Iceland (is)	2	4	**Sweden (se)**	12	12
Ireland (ie)	1	0	**Switzerland (ch)**	5	5
Latvia (lv)	0	1	**UK (uk)**	15	15

Table 1 Comparison of radar numbers in OIFS and NIMBUS databases.

Subsequently, we did not focus on an in-depth comparison of all the attributes of the datasets, but only on the occurrence of the radial wind parameter (VRAD) and the associated Nyquist velocity value. More information about general comparison of OIFS and NIMBUS radar data are in *Neštiak (2024).*

Available Nyquist velocity (NI)

Nyquist velocity, named after Harry Nyquist, is a key concept in Doppler radar systems. It represents the maximum unambiguous velocity that the radar can measure without aliasing. Aliasing occurs when higher velocities appear to be lower due to the limitations of the radar sampling rate. This happens because the radar samples the Doppler shift at discrete intervals determined by the pulse repetition frequency (PRF).

In radar meteorology and aviation (such as in remote control), accurately measuring velocities is crucial for tracking objects and understanding weather patterns. The Nyquist velocity helps ensure that these measurements are reliable and accurate, preventing errors that could lead to incorrect interpretations of the radar data (*Sireci, 2005).* The Nyquist velocity (*NI*) is determined by the radar's PRF. The relationship can be expressed by *Brown and Wood (2007)*:

$$NI = \frac{PRF \,.\, \lambda}{4}$$

where:

- λ is the wavelength of the radar signal.
- PRF is the pulse repetition frequency, i.e., the rate at which radar pulses are transmitted.

The factor of 4 comes from avoiding aliasing in both positive and negative velocity directions, effectively doubling the required sampling rate for accurate velocity measurement. In practice, choosing an appropriate Nyquist velocity depends on the specific requirements of the radar application. For instance, in meteorology, a higher Nyquist velocity is beneficial for observing fast-moving weather phenomena, while a lower Nyquist velocity might be more suitable for detecting slower-moving objects.

In the next step, data from NIMBUS were analysed and used. In the comparison of countries based on the condition of the Nyquist velocity value is shown with colour coding (Figure 1.1). Green represents data from countries whose datasets contained only NI values higher than 30 m/s. Red represents countries with values lower than 30 m/s. Yellow illustrates countries with both values higher and lower than 30 m/s. Countries marked in grey are missing either the value of NI (Czech Republic) or the entire dataset (Ireland, Romania, Serbia).

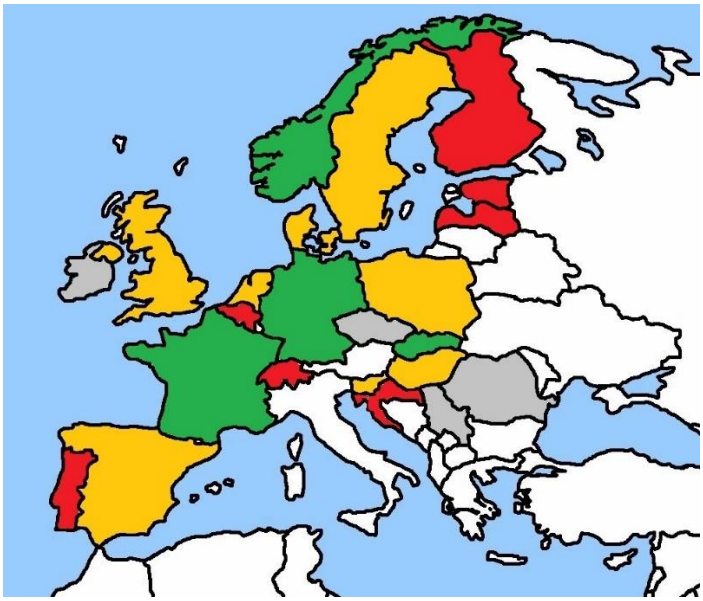

Figure 1.1 Colour code map of selected European countries based on the condition of the Nyquist velocity value.

Summary of data from selected countries

An overview of the available variables and the range of NI for the countries selected for this study from the randomly selected date of 1 April 2024 at 10 UTC is given in Table 2.

		Čatlošová (2020)	Petrovič (2024)			0 = not find
Country	**Radial winds**	**NI range [m/s]**	**NI range [m/s]**	**NIMBUS**	**Quantities with VRAD**	**Quantities with NI velocity limit set to 30 m/s**
Belgium (be)	available	<7>	<7>	<7>	0	0
Czech (cz)	available	0	0	0	VRADH+DBZH+TH	0
Croatia (hr)	available	17	<6,8,17>	<6,8,17>	VRADH+DBZH+TH	0
Denmark (dk)	available	<8,47>	<8,47>	<8,47>	VRADH+DBZH+TH+LDR +RHOHV+WRADH+PHDP +ZDR	VRADH+DBZH+TH+LDR+RHO HV+WRADH+PHDP+ZDR
Estonia (ee)	available	<8,13>	<8,13>	<8,13>	VRADH+DBZH+TH	0
Finland (fi)	available	8	<8>	<8>	VRADH+DBZH+TH	0
France (fr)	available	<59,62>	<59>	<59>	VRADH+DBZH+TH	VRADH+DBZH+TH
Germany (de)	available	<32>	<32>	<32>	VRADH+DBZH+TH	VRADH+DBZH+TH
Hungary (hu)	available	<8,47>	<8,47>	<8,47>	VRADH+DBZH	VRADH+DBZH
Iceland (is)	0	0	<30>	<30>	0	0
Ireland (ie)	0	0	<6>	0	0	0
Latvia	available	0	0	<7>	0	0
Malta (mt)	0	0	0	0	0	0
Netherlands (nl)	available	<6,80>	<6,32,48>	<6,32,48>	VRADH+DBZH+TH	VRADH+DBZH+TH
Norway (no)	available	<48,111>	<32,128>	<32,128>	VRADH+DBZH+TH	VRADH+DBZH+TH
Poland (pl)	0	7	<7,30,48>	<7,30,48>	VRADH	VRADH
Portugal (pt)	available	<6,50>	<6,8,16,60>	<6>	VRADH+DBZH+TH	0
Romania (ro)	0	8	0	0	0	0
Serbia (rs)	0	0	0	0	0	0
Slovakia (sk)	available	<40,64>	<40,64>	<40,64>	VRADH+DBZH+TH	VRADH+DBZH+TH
Slovenia (si)	available	<8,41>	<8,40>	<8,40>	VRADH+DBZH+TH	VRADH+DBZH+TH
Spain (es)	available	<3,48>	<3,48>	<3,48>	VRADH+DBZH	VRADH+DBZH
Sweden (se)	available	<24,40>	<24,40>	<24,40>	VRADH+DBZH+TH	VRADH+DBZH+TH
Switzerland (ch)	available	<8,12>	<8,12>	<8,12>	VRADH+DBZH+TH	0
UK (uk)	available	<4,48>	<4,48>	<4,48>	VRAD	VRAD

Table 2 Data overview of RC LACE countries with available radial winds, range of NI and quantities with VRAD. For comparison, the range of NI from the year 2020 (Čatlošová, 2020).

2. Preprocessing by HOOF

The HOOF Python tool was used for processing OPERA data. This tool ensures the homogenization of HDF5 data from OPERA to make them usable in BATOR. More information can be found in the documentation (HOOF manual and HOOF2 manual). Initially, we intended to use HOOF2; however, due to persistent issues with library versions in Python, it was not possible to use HOOF2. Initially, we compared datasets, but later the work focused exclusively on NIMBUS data.

Since the new NIMBUS data contain a new quality flag "eu.opera.odc.hac," which does not have mandatory *gain & offset* attributes, we used a workaround proposed by Alena's to overcome this. She proposed checking if the quality flag should be present in the dataset, and if so, copying the *gain & offset* from another group within the dataset and appending it to the new quality flag. This modification did not affect further calculations, as it only involved checking for the presence of the attribute, without further processing.

After processing the data with HOOFv1.9, warnings and errors were observed:

- Warning: A TH quantity from /dataset9/data1 has no matching DBZ, omitting the TH dataset.
- DBZ quantity in /dataset6/data1 does not have the required quality groups.
- Cannot sort file into measurements (probably no DBZ quantity is found).
- Attribute /dataset12/how/NI is required, but not found anywhere.

We selected datasets from the radars of Germany, Denmark, France, Hungary, the Netherlands, Poland, Slovenia, Slovakia, and the United Kingdom. The selection was based on attributes in the datasets (radial winds VRAD and Nyquist velocity were mandatory) and considering the computational domain of ALADIN/CHMI. The Michal Neštiak's results and findings were highly consistent with our own (Neštiak, 2024).

3. Bator

BATOR is a preprocessing tool used in the ALADIN Numerical Weather Prediction (NWP) model. Its primary function is to ingest and preprocess observational data, including radar, satellite, and in-situ measurements, to be used in the data assimilation system of ALADIN.

Key Functions of BATOR (Webpage HIRLAM.github.io):

- **Data Ingestion:** BATOR reads various observational data formats and converts them into a format that can be used by the ALADIN data assimilation system.
- **Quality Control:** The tool applies various quality control measures to ensure that the ingested data is accurate and reliable. This includes checks for consistency, validity, and redundancy.
- **Spatial and Temporal Matching:** BATOR aligns the observations spatially and temporally with the model grid and forecast time steps, ensuring the data is correctly positioned for assimilation.
- **Filtering and Cleaning:** The software performs filtering and cleaning operations on the data to remove noise and erroneous measurements. This step is crucial for radar data, which can be affected by various sources of noise and clutter.

3.1. Spatial Filtering of radial winds

One of the critical functions of BATOR when dealing with radar data is spatial filtering, which aims to clean and smooth the data before assimilation. This is necessary to remove spurious data points and ensure the quality of the observations (MONTMERLE and FACCANI, 2009).

Here's a breakdown of the process of the spatial filtering process in the BATOR system, specifically for Doppler wind data (denoted as 'DOPW'). The following sections (3.1.1 – 3.1.3) are interpretations of the source code subroutines (bator_decodhdf5_mod.F90, bator_util_mod.F90) :

Initial Cleaning:
The *bator_radar_wind_cleaner* subroutine is called with *filter* set to .TRUE. to perform an initial cleaning. This involves analysing the elevation data and identifying bad pixels or problematic elevations.

Filter Application:
If the data is not empty (i.e., there is valid wind data), the *bator_filter_radar* subroutine is applied to filter the data based on predefined criteria.

Final Cleaning:
A second call to *bator_radar_wind_cleaner* with filter set to .FALSE. is made to further clean the data at the pixel level.

3.1.1. Bator_radar_wind_cleaner (Wind cleaner I – Elevation check)

The subroutine *bator_radar_wind_cleaner* is responsible for eliminating bad pixels and handling elevations with frequency capture problems (CAF) for radial wind data. Here's a closer look at its operations:

1. **Initialization:**
   - The subroutine initializes necessary variables and allocates memory for arrays that will hold the data and error values. Is called with the filter parameter set to **.TRUE.** for initial cleaning at the elevation level.
2. **Data Extraction:**
   - It extracts radar data into a 2D array *PTAB* and a binary mask *NBWAG* to indicate valid wind observations.
   - It counts the initial number of valid observations and prints this information.
3. **Median Filtering:**
   - A median filtering process is applied to smooth the data and remove outliers. This involves calculating the median error and value within a specified window around each data point.
   - For each data point in the radar grid, the subroutine calculates the median error and value within a specified window (e.g., 5x5 grid) around the point.
   - Errors are computed as the difference between the central point and its neighbours, and these errors are sorted to find the median.
   - The median error (*MEAN_ERR*) and value (*MEAN_VAL*) are computed. If there are enough neighbouring points, these values are used to evaluate the quality of the central data point.

4. **Threshold Checking:**
    - The subroutine checks if the absolute difference between the observation and the median value exceeds a threshold. If it does, the data point is either removed or adjusted based on the filtering mode (elevation or pixel).
5. **Final Observation Count:**
    - The final number of valid observations is counted and printed, and if the ratio of remaining observations is too low, the entire elevation is rejected.

3.1.2. Bator_filter_radar (Median Filter)

The *bator_filter_radar* subroutine performs a median filter on the Doppler radial velocity data from radar observations. Here is a step-by-step explanation:

1. **Data Preparation:**
    - Loop through the radar data to populate NBWAG and PTAB_FILTRE.
    - Depending on the radar variable (*cvar*), set the appropriate offset *(ioff*).
2. **Median Filtering:**
    - For each pixel, identify a neighborhood window of size (2*DX+1) x (2*DY+1).
    - Count the valid data points within this window.
    - If the number of valid points exceeds the threshold (*COEFMEDIAN*), compute the median:
        - Sort the values in the window.
        - Find the median value and assign it to *PTAB_FILTRE*.
    - If not enough valid points are found, mark the pixel as invalid (*RABSO*).
3. **Update Radar Data:**
    - Loop through the data again to update the original radar data array (*ztwag*) with the filtered values.
    - Adjust the quality flag and observation count (*ztent*) accordingly.

3.1.3. Bator_radar_wind_cleaner (Wind cleaner II – Pixel check)

After median filtering, the cleaning process ensures that any remaining noise and errors are addressed. Here is how the cleaning process works in BATOR:

1. **Pixel-Level Cleaning:**
    - The subroutine *bator_radar_wind_cleaner* is called again with filter set to **.FALSE.** to perform detailed cleaning at the pixel level.
    - This step focuses on individual data points that may still contain errors after the median filtering.
2. **Threshold Application:**
    - For each data point, if the difference between the observation and the median value exceeds a higher threshold (e.g., 10 m/s), the data point is either removed or adjusted.
    - If the data point is an outlier, it may be replaced with the median value of its neighbours.
3. **Error Correction:**
    - Additional checks are performed to identify and correct significant errors. For instance, if the error between a data point and its neighbours exceeds 10 m/s, the point is flagged and corrected.

4. **Final Observation Count:**
    - The final number of valid observations is counted and compared to the initial count. If too few observations remain (e.g., less than 15% of the original data), the entire elevation might be rejected.
5. **Elevation Rejection:**
    - If an elevation contains too many errors or too few valid observations, it is rejected, and a message is printed indicating the rejection.

We selected two relatively random dates for our study: one during which we knew the atmospheric conditions were relatively calm (09.04.2024 at 11 UTC), and another during which a frontal system was passing through (01.04.2024 at 20 UTC). The data for these dates were recorded from the Slovak radar Malý Javorník. This selection is relevant to our analysis involving the wind cleaner and radar filter subroutines, as it allows us to compare the effectiveness of these subroutines under different atmospheric conditions. By doing so, we aim to understand how individual subroutines affect the resulting observations.

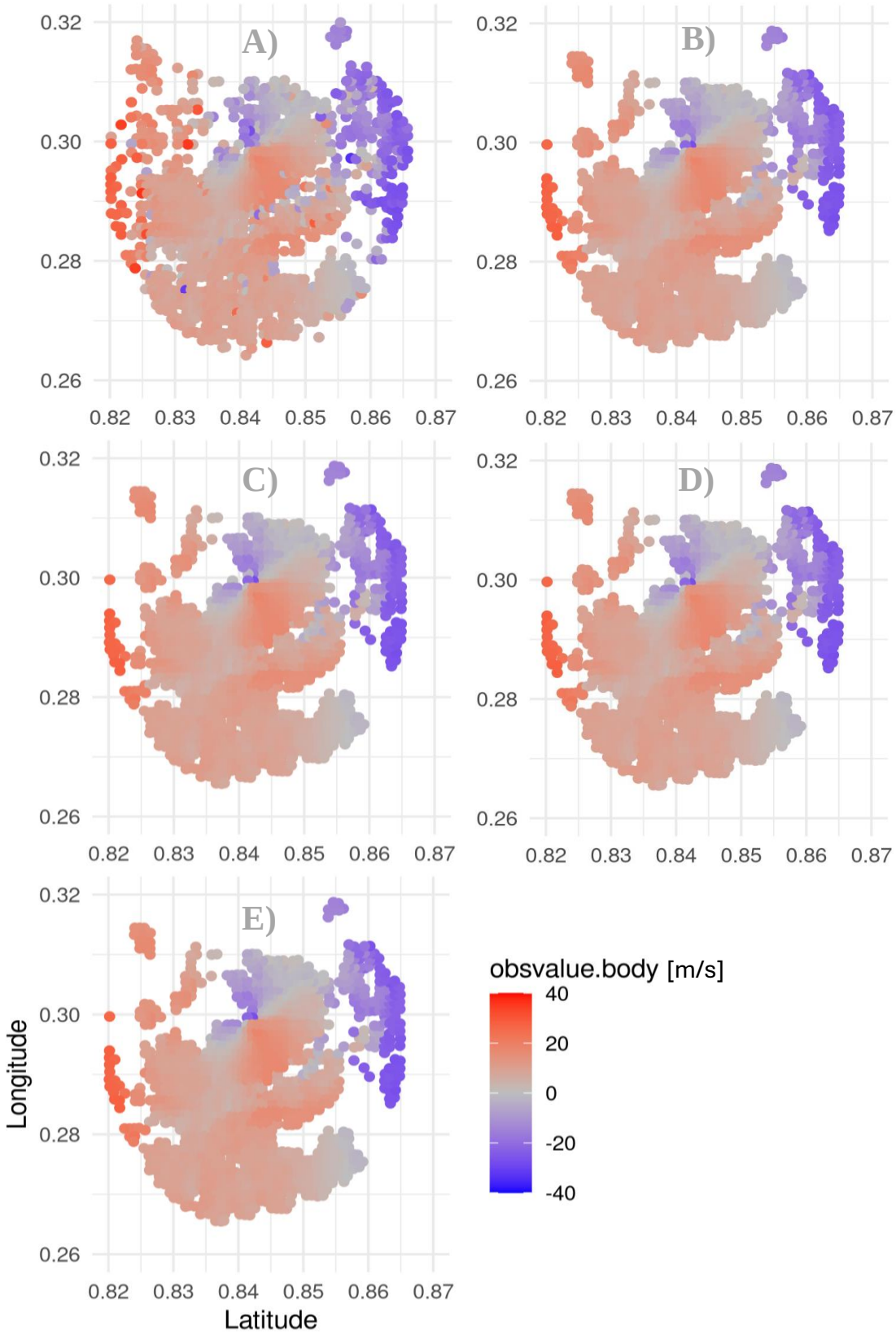

Figure 3.1 Visualisation of spatial distribution of observation after spatial filtering. Without spatial filtering output (A), only Median filter (B), Median filter + Pixel check (C), Elevation check + Pixel check (D) and whole spatial filtering = Elevation check + Median filter + Pixel check (E). Valid for randomly selected date 01.04.2024 at 20 UTC.

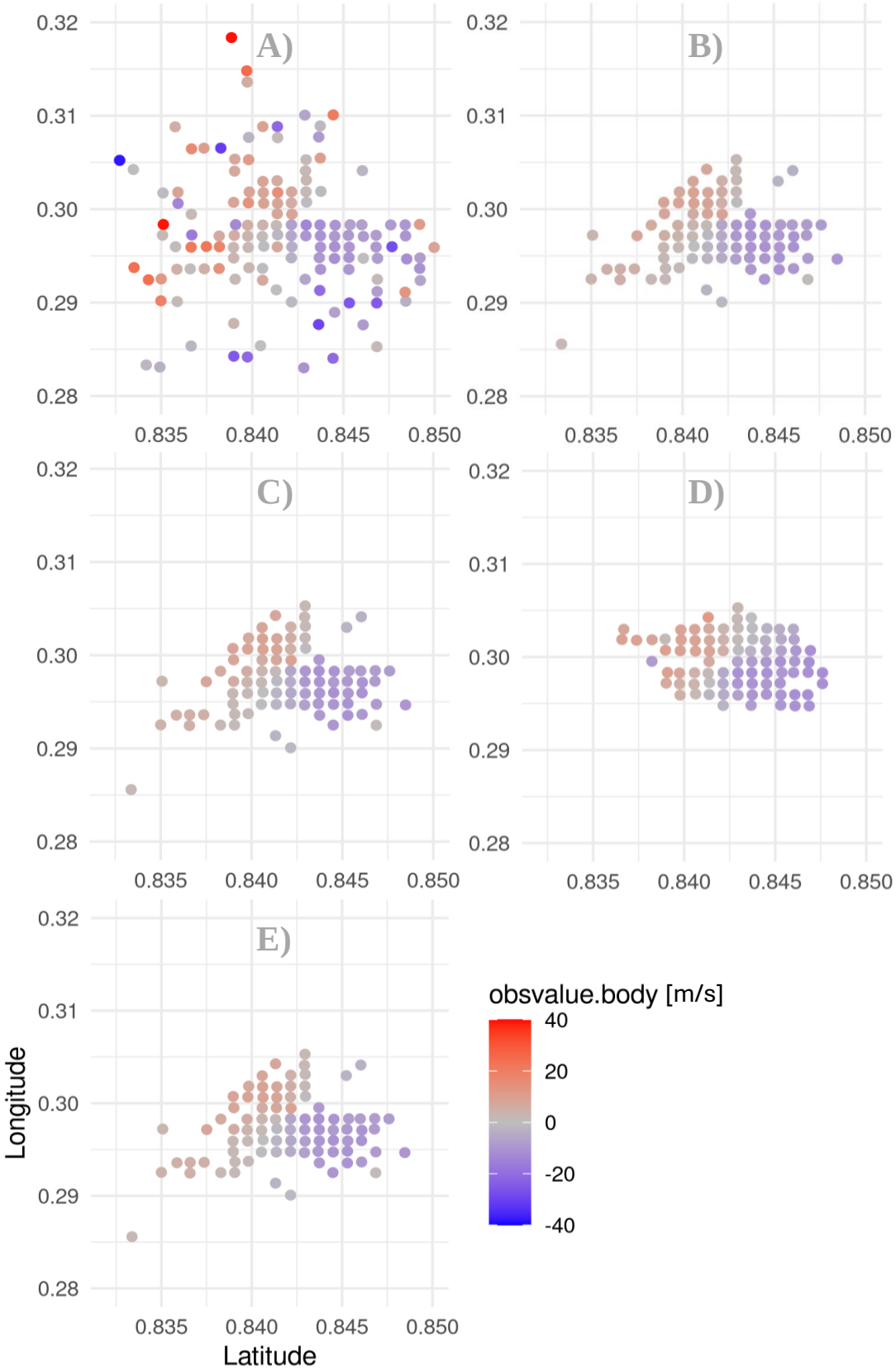

Figure 3.2 Visualisation of spatial distribution of observation after spatial filtering. Without spatial filtering output (A), only Median filter (B), Median filter + Pixel check (C), Elevation check + Pixel check (D) and whole spatial filtering (E). Valid for randomly selected date 09.04.2024 at 11 UTC.

3.2. Sub-sampling

Sub-sampling in radar data processing is a technique used to reduce the amount of data by selecting a subset of observations at regular intervals. This technique helps in managing data volume, improving computational efficiency, and potentially enhancing the quality of the resulting data analysis by removing redundant information.

Analysation of subroutine **bator_decodhdf5_mod.F90**. The *ifreq* is calculated as the ratio of the sampling distance (*HODIM%Sample*) to the radar resolution (*HODIM%Resolution*), rounded to the nearest integer. In the basic setting, the *ifreq* variable has a value of 5, from the above relationship. This means that every fifth pixel is kept and used in the next calculation. Subsequently, the value of the *ifreq* variable was changed to 1 and 10, which is shown in the following figure in comparison with the value of *ifreq* = 5.

As we mention, we tested three different sub-sampling intervals: 10 pixels, 5 pixels, and 1 pixel, with each interval representing the number of pixels between sampled observations. The basic setting for sub-sampling was every 5 pixels, which served as the reference point for our comparisons. The key metrics analysed were the number of observations left after sub-sampling and the effect on selected observations and data counts.

1. **Sub-sampling every 10th pixels**
   - The initial observation count was 86400.
   - After sub-sampling, 1798 observations were retained.
   - Out of these, 313 observations were selected, resulting in 4443 data points.

2. **Sub-sampling every 5th pixels**
   - The initial observation count was 86400.
   - After sub-sampling, 7236 observations were retained.
   - Out of these, 1257 observations were selected, resulting in 17617 data points.

3. **Sub-sampling every 1 pixel**
   - The initial observation count was 86400.
   - After sub-sampling, 75464 observations were retained.
   - Out of these, 25061 observations were selected, resulting in 340896 data points.

As we expected, decreasing the sub-sampling interval (i.e., reducing the number of pixels between sampled observations) resulted in a higher number of left observations. This increment in the number of observations naturally led to a greater volume of data points available for the assimilation process. The difference in observation counts and selected data points across the three sub-sampling settings demonstrate the trade-offs involved.

The sub-sampling experiment underscores the importance of balancing computational efficiency with the need for sufficient observational data in the assimilation process. While the basic setting of thinning every 5 pixels offers a moderate balance, the choice of interval should be guided by the specific requirements of the assimilation system and the computational resources available.

An open question remains whether the *ifreq* value can be set adaptively. The amount of retained observations significantly depends on atmospheric conditions, such as relatively calm weather or frontal systems. The purpose of this adaptive approach would be to use more observations during calm weather conditions and maintain the same or a relatively smaller amount during stormy weather (but no smaller amount than in basic set). This adaptive sub-sampling could potentially enhance the efficiency and accuracy of data assimilation by dynamically adjusting to varying weather patterns.

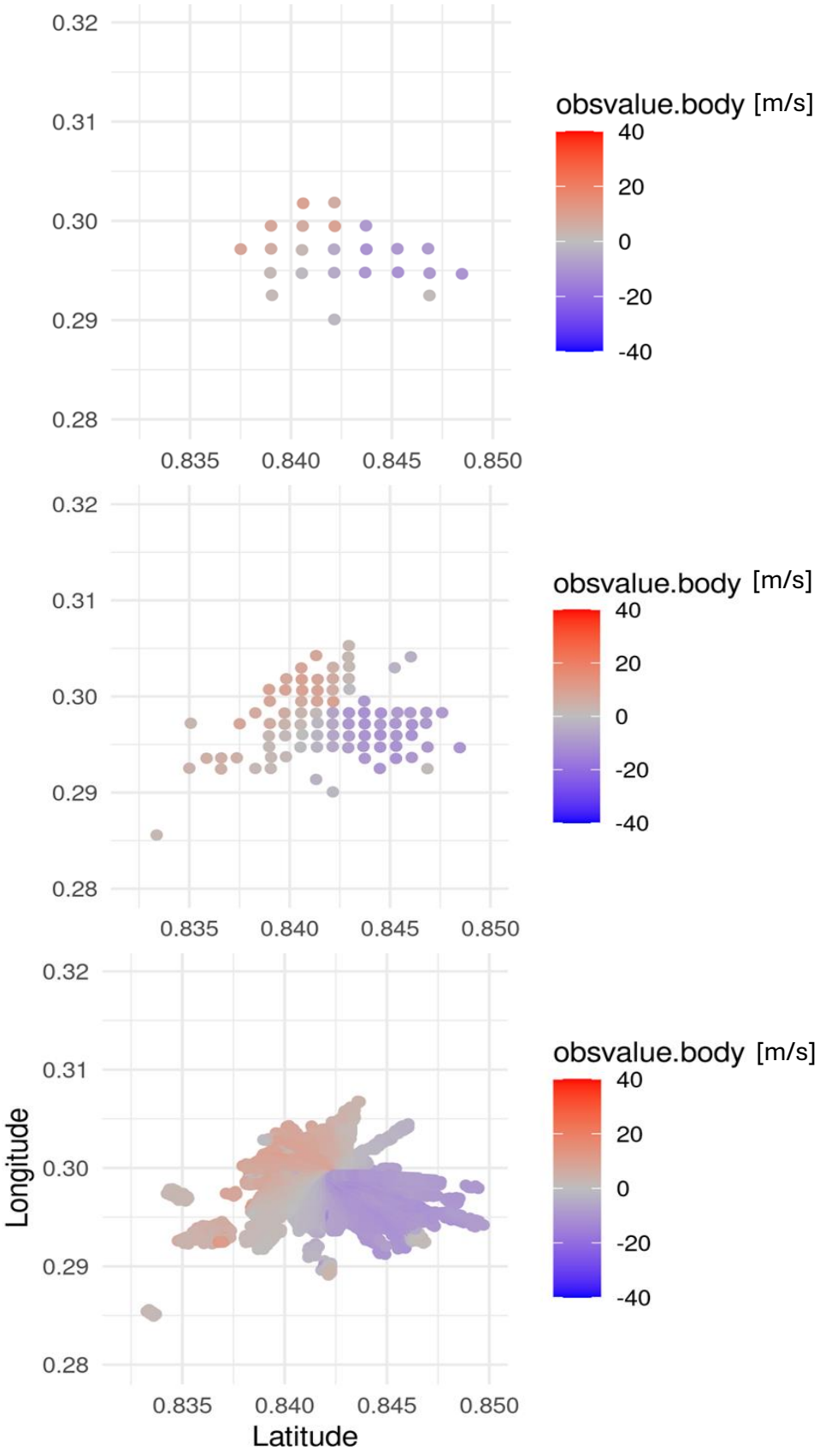

Figure 3.3 Visualisation of spatial distribution of observation after sub-sampling every 10th pixel (at top), every 5th pixel (in the middle) and every 1 pixel (at the bottom). Valid for randomly selected date 09.04.2024 at 11 UTC.

4. Passive data assimilation experiment

The preceding chapters provided an overview of the radial wind data and the process of their pre-processing. The next stage in the data assimilation chain is observation screening, which involves quality control, the removal of duplicated data, and a reduction in data resolution (thinning).

One way to assess the quality of new observations is to compare them with the NWP background. Passive data assimilation experiments enable the computation of differences between observations and model forecasts (OMG) without affecting the analysis.

A passive experiment with radar observations was conducted using the setup version cy46t1mp_op3 at CHMI (Czech Hydrometeorological Institute) on areas covering central, southern and part of western Europe. It was launched for the period 01.04.2024 at 00 UTC to 15.04.2024 at 18 UTC. We run a 3-hour assimilation cycle with selected radars. The Nimbus data were homogenised, as is written in section 2., and then were read by BATOR. We used available data from all radars having NI > 30 m/s from the selected countries (number of used radars):**"de (17)", "dk (4)", "fr (23)", "hu (5)", "nl (2)", "pl (4)", "si (2)", "sk (4)", "uk (11)** ".

Firstly we checked difference between amount of all observations and active observations of radar doppler wind (DOW, varno = 195). Number of observations are visualised in Figures 4.1 – 4.3. The plot visualizes the distribution of observations across different stations at same time periods. This helps identify stations with sufficient data coverage for further analysis. By comparing "Active Data" and "All Data" lines, one can assess data completeness and potentially identify gaps or inconsistencies. Also unusual spikes or drops in observation counts might indicate potential issues or outliers that require further investigation. Analysed spikes were referring to precipitation in the atmosphere. However, some drops (Figure 4.1– FRTRA after 11 Apr, Figure 4.3– PLGDA after 6 Apr) represent issues with datasets. The issues were mostly due to missing whole datasets or missing clue part of datasets.

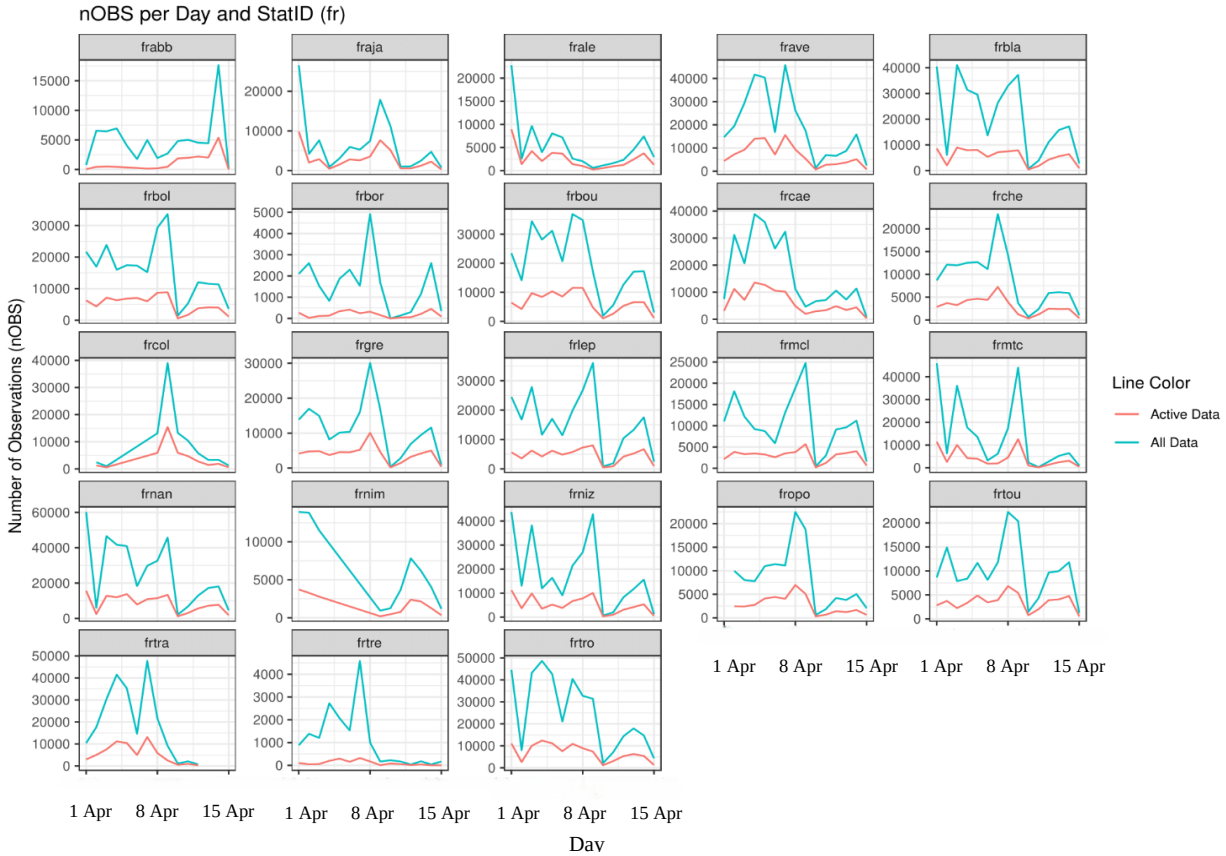

Figure 4.1 Visualization of the total and active Doppler wind observations from French radars over a two-week period.

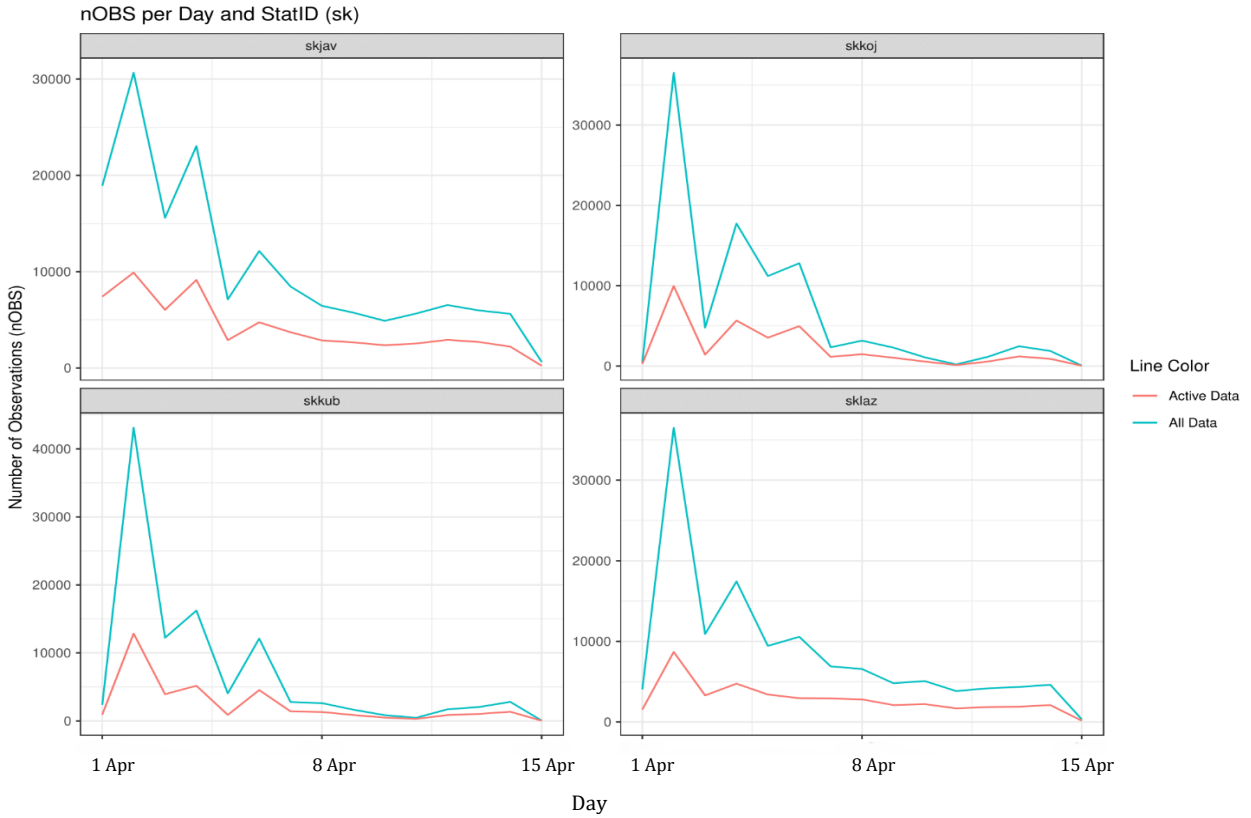

Figure 4.2 Visualization of the total and active Doppler wind observations from Slovak radars over a two-week period.

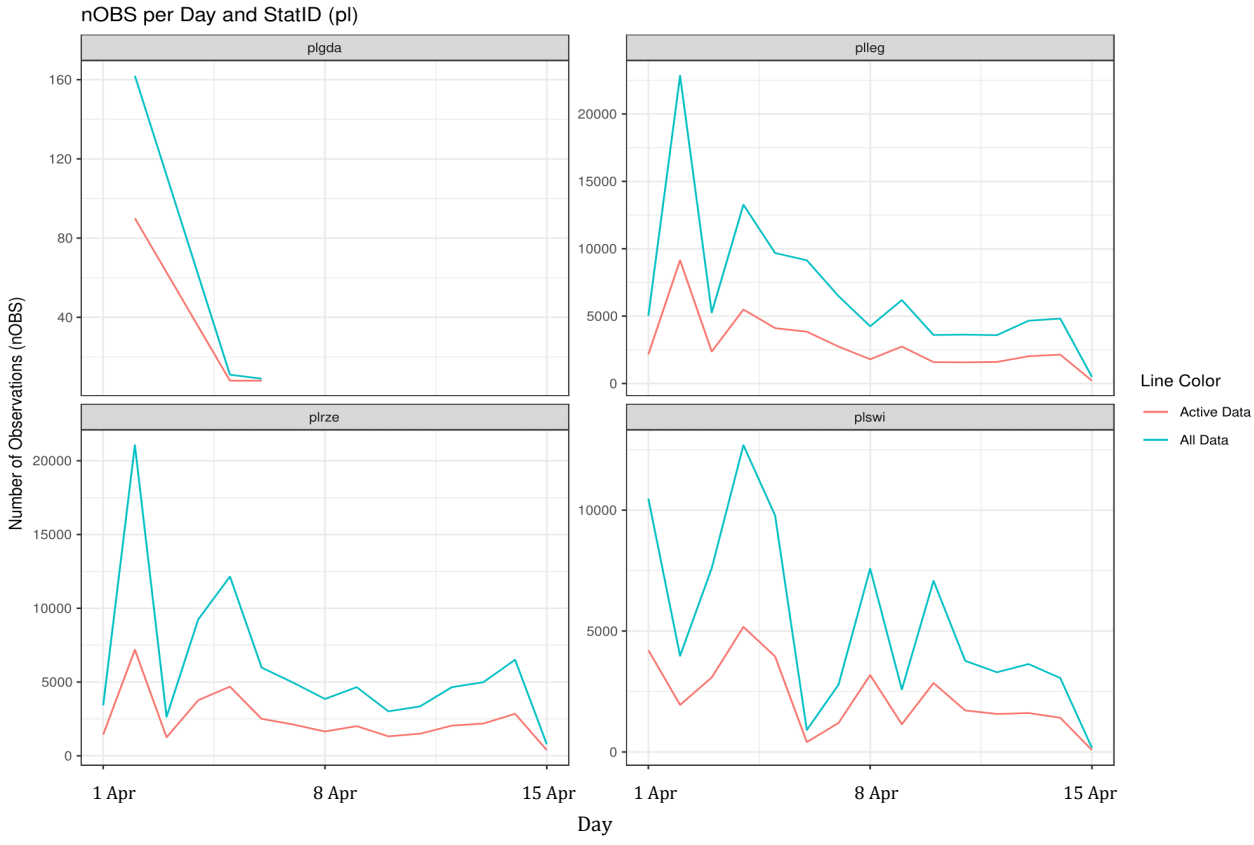

Figure 4.3 Visualization of the total and active Doppler wind observations from Polish radars over a two-week period.

4.1. OMG histograms

The provided histogram shows the distribution of differences between observations and model forecasts of Doppler velocity (Figure 4.4). The data is from a two-week period and includes both active and all data from selected countries. The mean difference is 0.24 m/s and the standard deviation is 5.52 m/s.

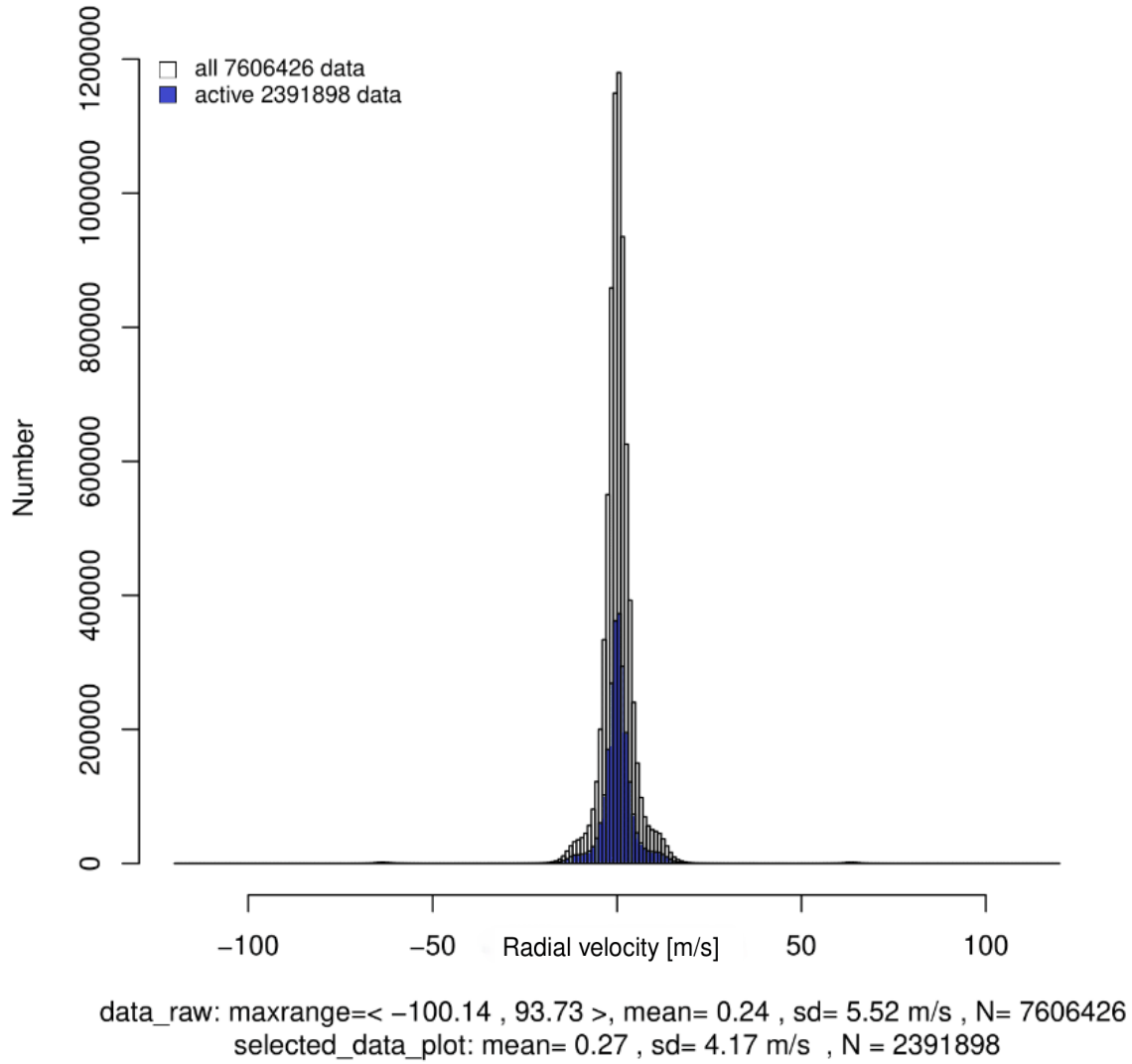

Figure 4.4 Histogram of observations-minus-guess departures, combining data from all selected countries for both active and all data over a two-week period.

For someone it can be seen, that histogram is not purely Gaussian, but very slightly bimodal and has with a peak near 0 m/s and a smaller peak near around10 m/s. This can suggest that there are sources of error.

The histogram for SKJAV (Figure 4.5) shows the Gaussian distribution (mostly around the central peak). The mean difference between observations and model forecasts is -0.46 m/s, slightly negative, suggesting a tendency for the model to overestimate Doppler velocities. The standard deviation of 7.01 m/s indicates a relatively high variability in the differences. The selected data (active) shows a similar distribution but with a slightly smaller mean -0.25 m/s and standard deviation 5.97 m/s.

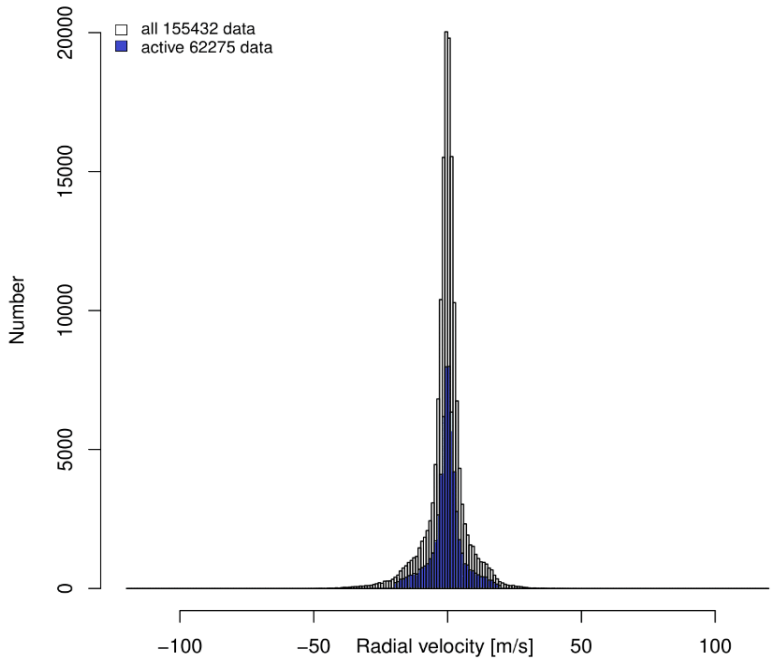

Figure 4.5 Histogram of OMG departures for Slovak radar Malý Javorník.

The histogram for DKVIR (Figure 4.6) shows a clear bimodal distribution, indicating two distinct groups of data. This suggests the presence of multiple error sources. Someone would say it could be caused by the weather regime, but a similar distribution was presented in Čatlošová (2020). The mean difference between observations and model forecasts is -2.92 m/s, indicating a significant negative bias in the model forecasts. The standard deviation of 26.32 m/s suggests a very high variability in the differences, possibly due to extreme weather events or other factors. The selected data has a mean of 0.85 m/s, indicating a positive bias. The standard deviation is reduced to 12.19 m/s, suggesting that the selection criteria effectively removed some outliers.

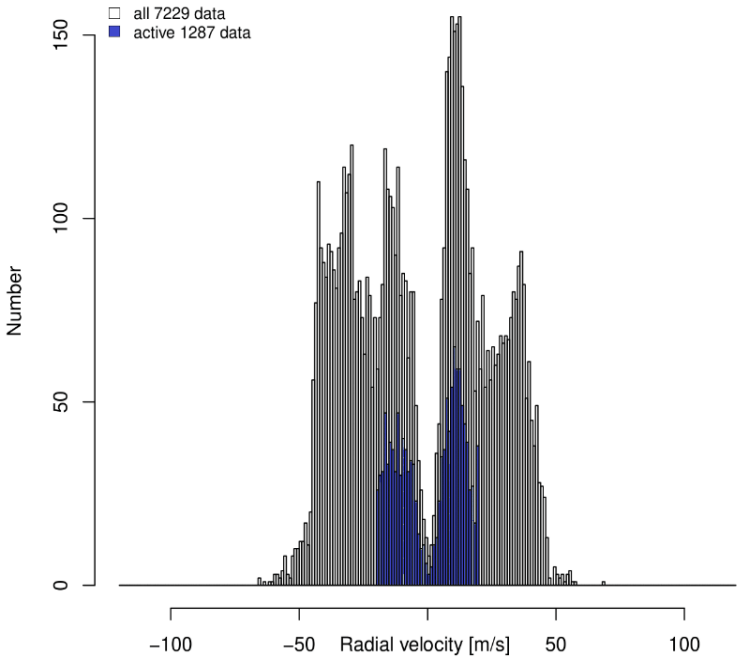

Figure 4.6 Histogram of OMG departures for Danish radar at Virring

It´s important to notice that a bimodal distribution was observed only in data from Danish radar at Virring, others Danish radars have the Gaussian distribution of departures.

In some cases, we observed two peaks far from the central peak (Figure 4.7, left one), approximately at the level of departures with a value of +- 60 m/s. This feature was observed in several datasets for all data from German radars. However, it was not observed for active data. We noticed another feature on some data from French radars (Figure 4.7, right one), where two smaller peaks appeared near the central peak. Similar feature was on the first histogram of departures.

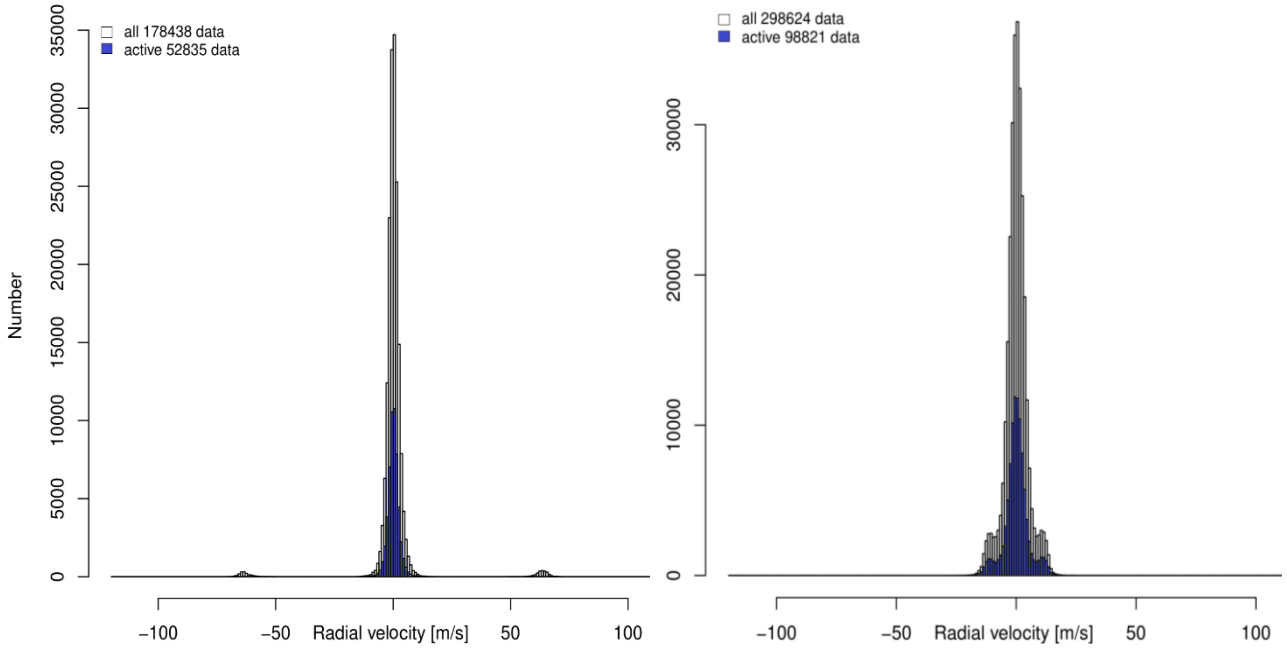

Figure 4.7 Histogram of OMG departures for German radar at Essen (left) and French radar at Bourges (right).

4.2. Rejection limit

The first-guess check step is crucial in determining the inclusion of an observation in data assimilation by comparing the observation with the first-guess or model background field. If the observations deviate significantly from the model equivalent, they are discarded.

To investigate the first-guess check limits, the Anderson and Järvinen (1999) technique was employed to examine the histograms and inverse histograms of background.

The Figure 4.8 (right) shows the transformed departures for active data of radar velocity from our experiment. The left panel shows the transformed histogram of observation departures for the background, which have been transformed according to the following equation:

$$f = \sqrt{-2 * \ln \frac{f}{\max(f)}}$$

where f is the number of data in each bin of the histogram. The slope of the points, indicated by red line, defines the standard deviation of the Gaussian curve that is represented in Figure 4.8 (right). 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.

From visual comparison the rejection limit can be set at a certain distance beyond the point where the values of function *f* start to deviate from the straight red lines. As seen in Figure 4.8 (right), the first-guess departures follow a Gaussian behaviour up to 5-6 m/s, but for higher values they form tails, indicating a non-Gaussian distribution. This is believed to be one of the reasons why the assimilation of radial velocity had a detrimental impact on the forecasts (Ridal, 2023).

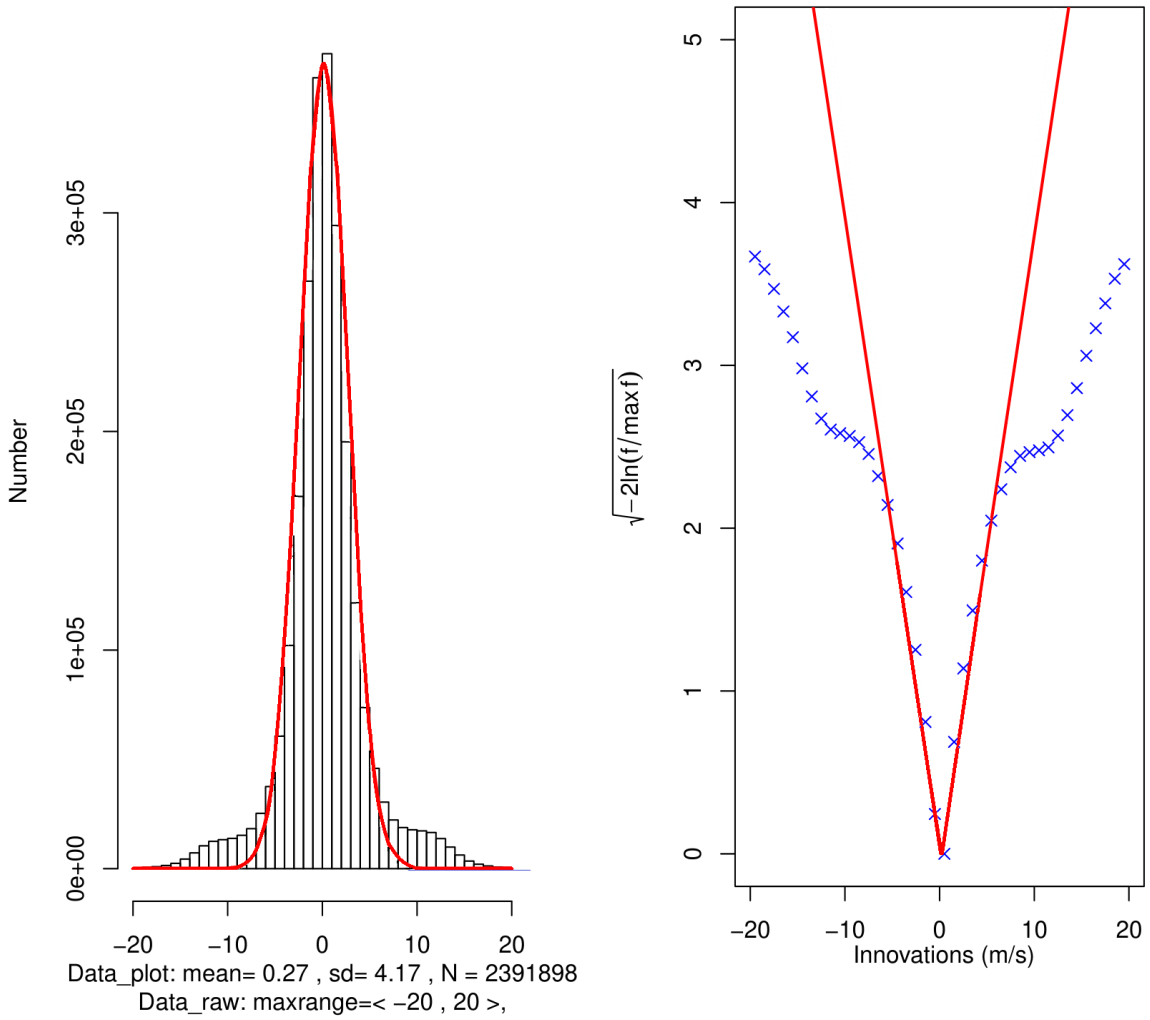

Figure 4.8 Histogram of OMG departures of active data (left) and transformed histogram (right).

For future work, it would be interesting to set a rejection limit (threshold) at 5-6 m/s, based on the findings above. Beyond this threshold, the data exhibit non-Gaussian tails, which likely contribute to the deterioration in forecast quality when assimilating radial velocities.

5. Conclusion

During my internship at CHMI, I focused on the assimilation and validation of radar radial wind data to improve weather forecasts within the ALADIN model. The main tasks involved comparing the OIFS dataset with the new NIMBUS dataset, analysing the differences between them, and applying spatial filtering and passive assimilation methods to these data.

The initial analysis focused on comparing the OIFS and NIMBUS datasets. We found discrepancies in both datasets, with missing stations in OIFS and NIMBUS, and varying radar counts per country. We did examine the radial wind parameter (VRAD) and Nyquist velocity (NI) values. The analysis used colour coding map to visualize the distribution of NI values across different countries, highlighting those with sufficient or insufficient coverage.

Datasets were pre-processed by HOOFv1.9. Homogenising the NIMBUS data had problems with incomplete quality flags (the problematic one "eu.opera.odc.hac"). This was solved with Alena´s "hack" to copy the *gain & offset.*

We analysed radar data filtering to remove unwanted artefacts and noise. This step proved crucial for enhancing data quality before assimilation.

The results of passive assimilation demonstrated a promising quality of radar radial winds in comparison to the NWP model. However, further efforts are required to refine the quality control of the data (first guess check). The transformed histogram provides a valuable tool for visually comparing radial wind departures with the theoretical distribution (in our case, the Gaussian). Additionally, it can serve as an approximation of the threshold.

Further experiments should focus on active assimilation and improving the model's predictive capabilities, with an emphasis on verifying the practical benefits for meteorological forecasts.

Acknowledgement

I would like to express my sincere gratitude to Alena Trojáková for her expert advice, valuable ideas, and for creating a very pleasant working environment at the Czech Hydrometeorological Institute (CHMI) under the auspices of the RC LACE. I would also like to thank Antonín Bučánek and the staff of the Department of Numerical Weather Prediction at CHMI for their contributions.

References

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

BROWN, R. A. – 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. Retrieved 16.4.2024.

ČATLOŠOVÁ, K. 2020. Assimilation of radial velocity from radars. *RC LACE 2020*. Available: https://www.rclace.eu/media/files/Data_Assimilation/2020/repStay_KCatlosova_RadialWindAssimilaton_2020_07.pdf . Retrieved 25.3.2024.

EUMETNET, 2024. OPERA Available: https://www.eumetnet.eu/activities/observations-programme/current-activities/opera/. Retrieved 29.8.2024.

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

NEŠTIAK, M. 2024. Testing of radar data from the new OPERA NIMBUS production line. *RC LACE 2024.* Available: https://www.rclace.eu/media/files/Data_Assimilation/2024/repStay_MNestiak_Nimbus_CHMI_2024.pdf . Retrieved 21.5.2024.

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 . Retrieved 16.4.2024.

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 . Retrieved 22.3.2024.

Source code of Bator. Version 46t1mp/Odb/pandor/module/*subroutines*. Retrieved April 2024.

Webpage HIRLAM.github.io**.** Harmonie wiki: Bator. *ODB creation: Bator.* From: https://hirlam.github.io/HarmonieSystemDocumentation/dev/Observations/Bator/# . [Online] Retrieved 10.7.2024.