

OPERA radar data use in AROME-France model

Maud Martet

2019

Contents

| | |
|---|-----------|
| Context | 2 |
| 1 OPERA radar data | 2 |
| 1.1 Characteristics of the data | 2 |
| 1.2 Quality flags | 2 |
| 2 Observations treatment in BATOR | 3 |
| 2.1 Standardization of the data | 3 |
| 2.2 Non-rainy observation and <i>undetected</i> value | 3 |
| 2.3 Radar minimum detectable signal calculation | 3 |
| 2.4 Selection of the data | 4 |
| 3 Results | 7 |
| 3.1 AROME-France configuration | 7 |
| 3.2 Systematic monitoring | 8 |
| 3.3 OPERA radar data assimilation: one month experiment | 9 |
| 3.3.1 Period of interest | 9 |
| 3.3.2 "Score indicateur" | 9 |
| 3.3.3 Comparison with ground stations over France | 10 |
| 3.3.4 Comparison with European ground stations | 10 |
| Conclusion | 12 |
| References | 13 |
| A Reading ODIM files in BATOR, available from cy43 | 14 |

Context

One goal of the OPERA program is to collect radar volume data and distribute quality flagged volume data to modellers. Data are available in the same format: the OPERA Data Information Model (ODIM). Odyssey (OPERA Data Center) runs operationally to collect and add quality flags to national radar data.

Radar data from non Météo France national weather service are usable in AROME-France and will be included for systematic monitoring this year. Figure 1 shows the positions of the radars included in the AROME-France domain.

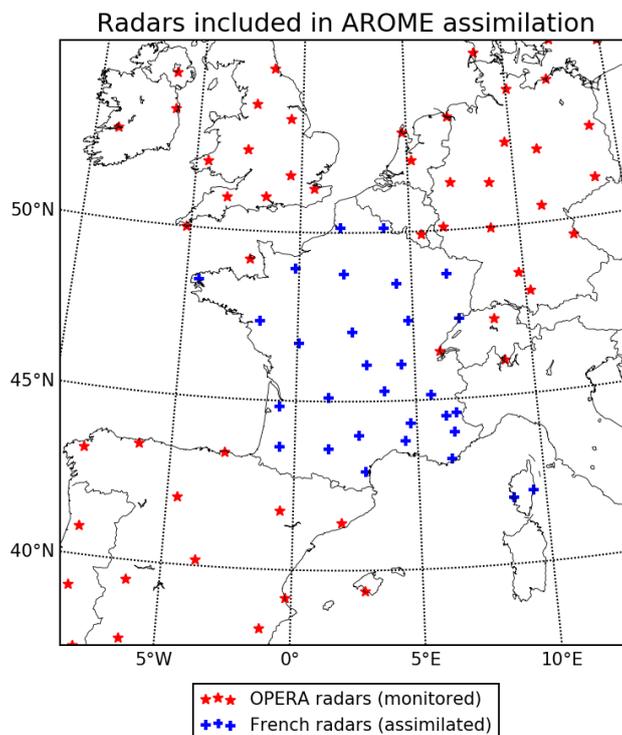


Figure 1: Radar included in the AROME-France domain.

1 OPERA radar data

1.1 Characteristics of the data

OPERA provides radar data from all the participating countries on an ftp server (OIFS, OPERA Internet File System). For each radar, all PPIs (Plan Position Indicator) from the last 15 minutes are gathered in a volume file, every 15 minutes. In the ODIM format, no specifications are made in order to harmonize elevations or horizontal resolution of the radar data. It is necessary to take this into account in the treatment of the data in BATOR. Some metadata are not mandatory in ODIM format, like beamwidth or radar constant. Actions are made with OPERA team in order to change this practice in the ODIM 2.3 version.

1.2 Quality flags

OPERA paradigm is (until 4th phase) to collect "raw" radar data and add quality flags in order to get homogeneous information all over Europe. So, a centralized process is run to add quality flags and to correct data from various observation errors (removal of non

meteorological echoes, correction of beam blockage for instance).

In the current Odyssey version, 3 treatments are applied:

- satellite filter to remove non meteorological echoes using Precipitating Clouds product from SAF-NWC (PGE04),
- BroPo from Baltrad toolbox to remove non meteorological echoes also in cloudy conditions,
- BeamB from Baltrad toolbox to correct data from beam blockage attenuation.

Each module adds a quality index and a total quality index gathers them taking the minimal value, as described in [Saltikoff et al., 2019].

2 Observations treatment in BATOR

2.1 Standardization of the data

In AROME-France, 1D+3D-Var assimilation method is implemented for reflectivity radar data. This method implies profiles of reflectivity to be used in a Bayesian method as explained in [Wattrelot et al., 2014]. However, in OPERA data, no specifications are given on horizontal or vertical resolutions. Each radar can have different elevations (numbers and angles) and each elevation can have a different number of azimuths or gates. So, in order to be able to consider vertical profiles of reflectivity, a choice has been made in BATOR: for each radar, the elevations with the number of azimuths with the two most "populated" are used, the others are not taken into account. A technical description of the method used in BATOR, written by Frank Guillaume, is available in appendix A.

Concerning radial velocity, another constraint is that no quality index is calculated on this parameter as on reflectivity. In order to be able to discriminate non-meteorological echoes, the quality index calculated for reflectivity is also used for radial velocity, when it is possible, meaning when reflectivity and radial velocity are observed during the same PPI or if a common elevation is used for reflectivity and radial velocity (even if the time of observation is not strictly equal). This approximation is not ideal but as we want to be sure not to assimilate radial velocity from non-meteorological echoes, it is the best compromise.

2.2 Non-rainy observation and *undetected* value

As explained [Saltikoff et al., 2019], in OPERA radar data, a common definition for dry pixels was needed and two terms were defined: *nodata* is used to describe that the pixel is out of range or in a blanked sector, *undetected* means that the received radar signal is at or below noise level. A problem occurs when ground clutters are removed and reflectivity value is set to *undetected*. In this case, there is a risk to assimilate data as "no rain" whereas we are not sure there is no rain. In order to discriminate dry area from ground clutter, it has been asked to NMS to send corrected reflectivity (DBZH) without ground clutter and uncorrected reflectivity (TH) with ground clutter. A comparison between DBZH and TH is performed: if DBZH value is *undetected*, when TH is also *undetected*, we are sure it is a non rainy pixel otherwise, reflectivity has been removed and there is a doubt whether the pixel is rainy or not. In order to assimilate non-rainy observation, DBZH and TH must be considered.

2.3 Radar minimum detectable signal calculation

In order to assimilate "non-rainy" observations, as explained in [Wattrelot et al., 2014], it is important to know the minimum detectable signal observable by each radar. When this information is not available in metadata, an approximation is made using the minimum value of reflectivity observed in the PPI. If this value is available, the sensitivity threshold

is calculated, function of the distance, in order to assimilate the "non-rainy" observations and to be able to dry the model where no rain is observed.

2.4 Selection of the data

In order to correctly assimilate radar data, only data with good quality index is taken into account. Quality indexes are available with the data, as explained in part 1.2. In order to determine a threshold that define "good quality", radar images have been studied. Two examples are shown below from 20/05/2019 12h00 UTC situation:

- Chenies radar (figure 2). Best radar data are already sent to OPERA from MetOffice, as it is specified from OPERA 5.
- Cerceda radar (figure 3). AEMET sends radar data as it was asked during OPERA 4: reflectivity with only elimination of ground clutters.

For each radar, PPIs from a low elevation are shown: the reflectivity as it is sent by NMS, the reflectivity as it is available in the files available from the OIFS, the total quality index and filtered reflectivity using various threshold on QI (0.6, 0.7 and 0.8).

Concerning Chenies radar, some non-meteorological echoes are visible in the vicinity of the radar even in MetOffice sends the best reflectivity. Odyssey toolbox allows to remove these echoes (figure 2d). Using a threshold on quality index to remove data which are not "good" has not a big impact on the rainy area located East of the radar. Nevertheless, on the light rainy spot South of the radar, using 0.8 as threshold removes this information.

Concerning Cerceda radar, some non-meteorological echoes are visible in the reflectivity sent by AEMET. This is expected as AEMET has not applied OPERA 5 specifications yet. Odyssey toolbox allows to remove lots of non-meteorological echoes but also removes "true" rain due to the too aggressive satellite filter. Using a threshold on quality index increase this phenomenon but gives more confidence on the fact that the reflectivity kept is really rain. On this particular example, 0.6 and 0.7 seem to be acceptable thresholds but as we want to be sure not to use non-meteorological echoes, the highest one is preferred.

After such a qualitative study, the threshold has been set to 0.7. Concerning "non-rainy" assimilation, the minimum detectable signal (MDS) is calculated from the metadata or dynamically as explained in the previous section. In order to efficiently assimilate "non-rainy" observations, we chose to use this information only where the MDS is under 0 dBZ.

Concerning radial velocity, we chose not to consider radial velocity when Nyquist velocity (NI) is under a threshold set to 30 m/s. This value is a compromise between avoiding aliased radial velocities and keeping data from some radars (lots of radial velocity PPIs have smaller Nyquist velocities). If the Nyquist velocity is not present in the "/dataset/how" group or in the top "/how" group, the radial velocity of this dataset is not included in the assimilation system.

In BATOR, a first sampling is applied to keep data every 5 km. This distance is sufficient given the thinning applied in the screening.

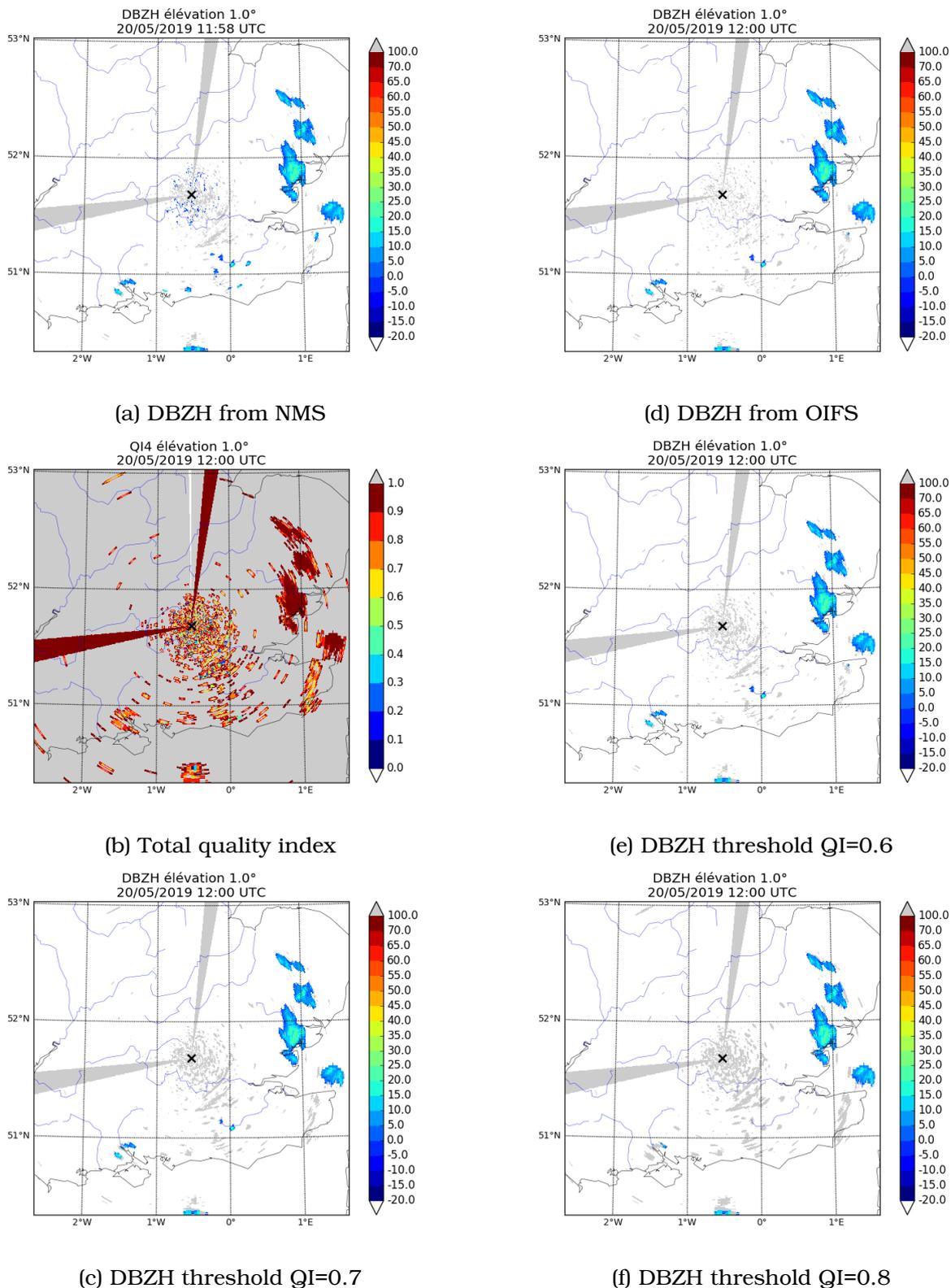
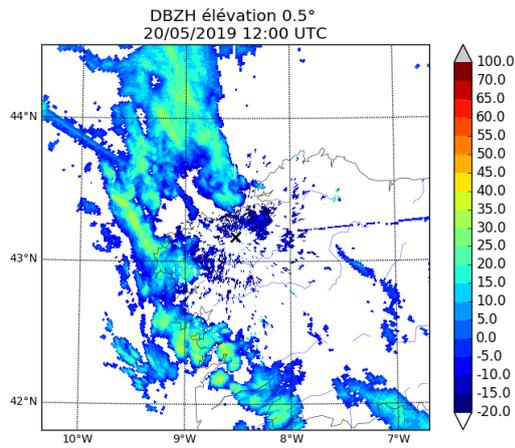
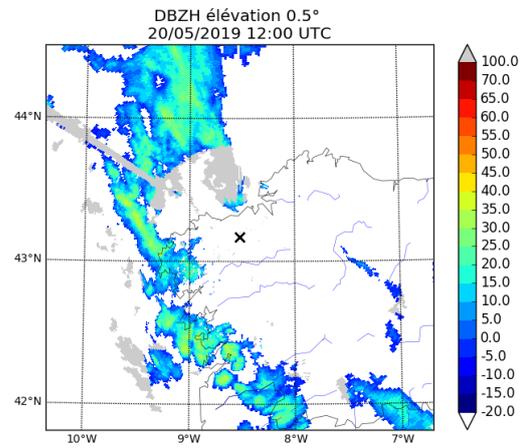


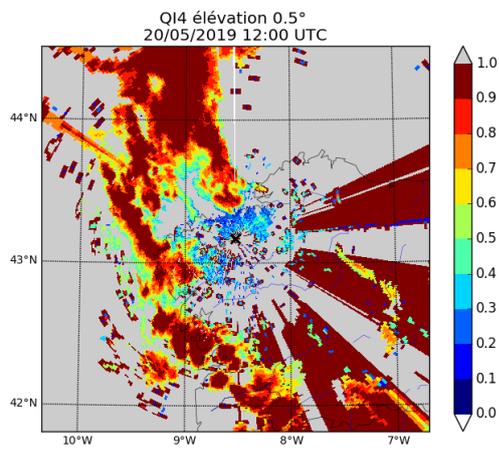
Figure 2: PPIs from Chénies (United Kingdom) radar, 20/05/2019 12h00 UTC, elevation 1.0°.



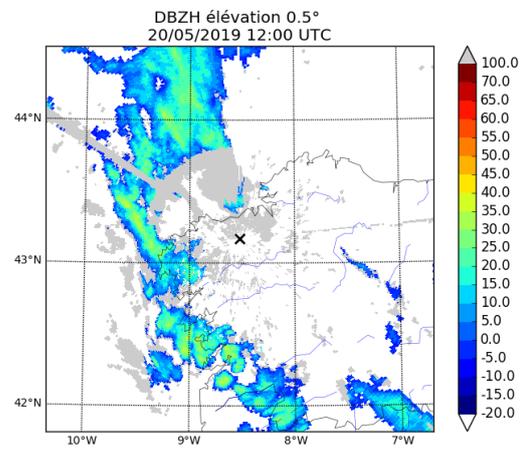
(a) DBZH from NMS



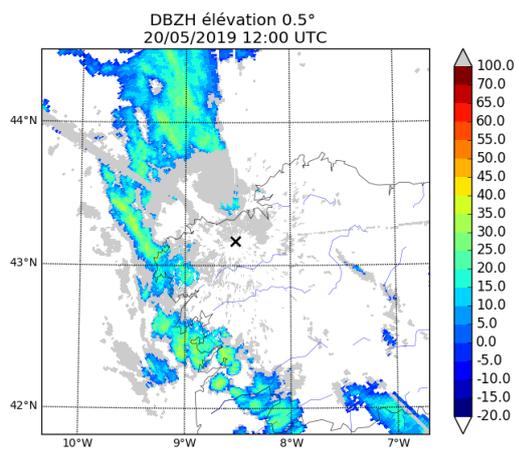
(d) DBZH from OIFS



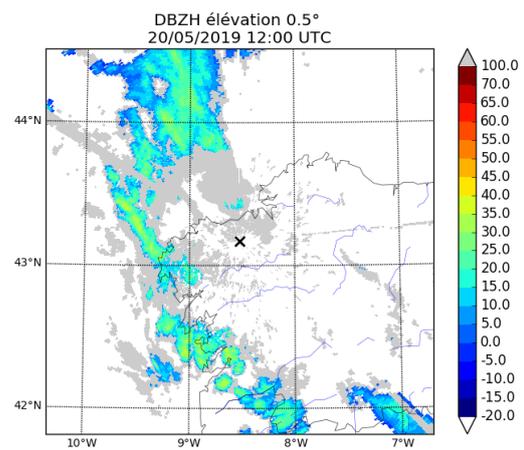
(b) Total quality index



(e) DBZH threshold QI=0.6



(c) DBZH threshold QI=0.7



(f) DBZH threshold QI=0.8

Figure 3: PPIs from Cerceda (Spain) radar, 20/05/2019 12h00 UTC, elevation 0.5°.

3 Results

3.1 AROME-France configuration

AROME-France configuration consists of a 1-hour assimilation cycle, an horizontal resolution of 1.3 km and 90 vertical levels as described in [Brousseau et al., 2016]. Radar data are thinned at 8 km resolution as described in [Wattrelot, 2016].

In this study, the Météo France radar data product is used for French radars and the OPERA product is used for the 62 non-Météo France radars included in AROME-France domain. As explained in part 2.4, reflectivity is used when DBZH, TH and QI are available, radial velocity is used when Nyquist velocity is in the metadata and greater then 30 m/s. Unfortunately, metadata and radar data are not available in all the datasets. Figure 4 represents the data (reflectivity and radial velocity) used by country. Table 1 gives the explanation why data are used or not by country. We keep on asking OPERA to make more data and metadata available but this is rather a slow process.

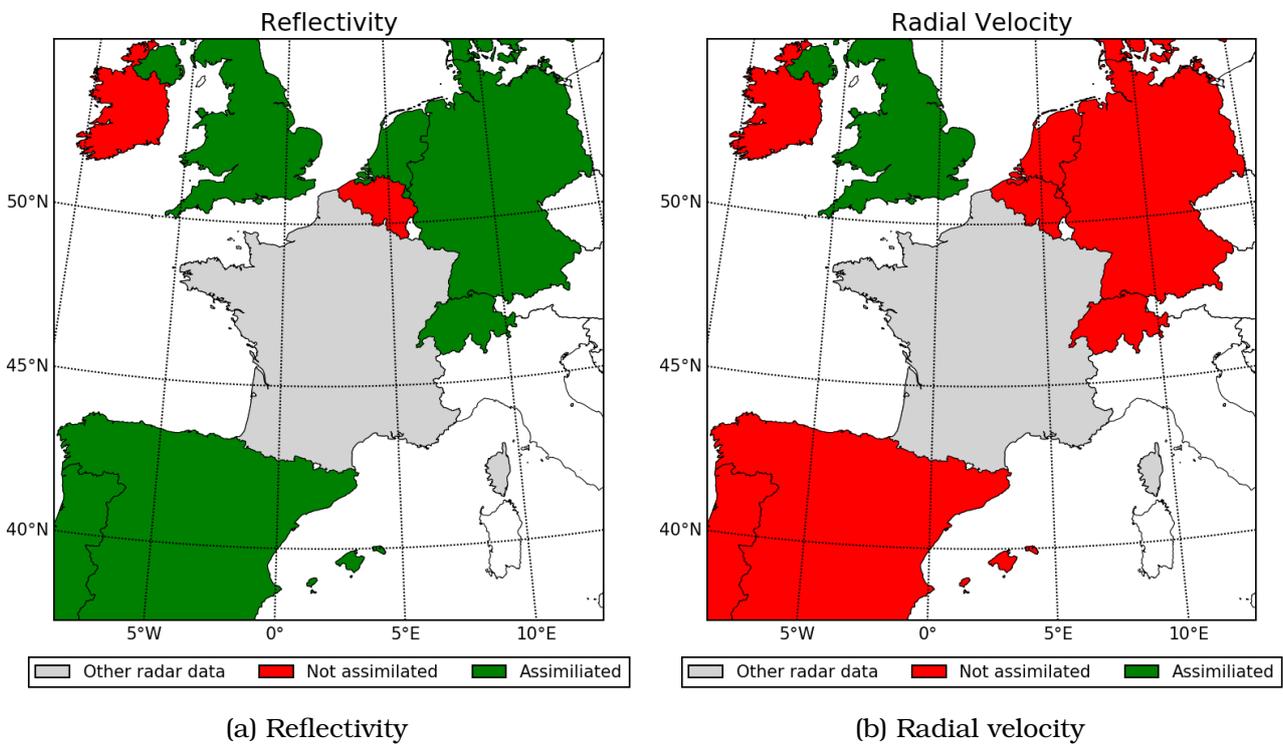


Figure 4: Data used function of data availability by country.

| | | | | | |
|-----------------|----------|------------|-------------|----------------|-----------------|
| | Belgium | Denmark | Germany | Ireland | Netherlands |
| Reflectivity | no TH | OK | OK | no TH | ok |
| Radial velocity | no VRAD | NI too low | no NI | no VRAD | wrong NI format |
| | Portugal | Spain | Switzerland | United Kingdom | |
| Reflectivity | ok | ok | ok | ok | |
| Radial velocity | no VRAD | NI too low | NI too low | ok | |

Table 1: OPERA radar data usage in AROME-France.

3.2 Systematic monitoring

Since 2019 January 23rd, OPERA radar data are included in AROME-France for systematic monitoring. Histograms of innovations allow us to determine the radar data quality. Figure 5 represents the histograms of innovations of OPERA radars data integrated in AROME-France on a 10 day period (from 2019/01/24 to 2019/02/02). All data available in the screening are considered. OPERA radars data are not assimilated in this experiment. For the two parameters, the innovation distribution follows an unbiased Gaussian law. Figure 6 represents the innovations for Météo France radar data over the same period. Innovations are comparable with OPERA radar data. The main difference is the number of radial velocity observations: less radial velocity data are considered, as explained above.

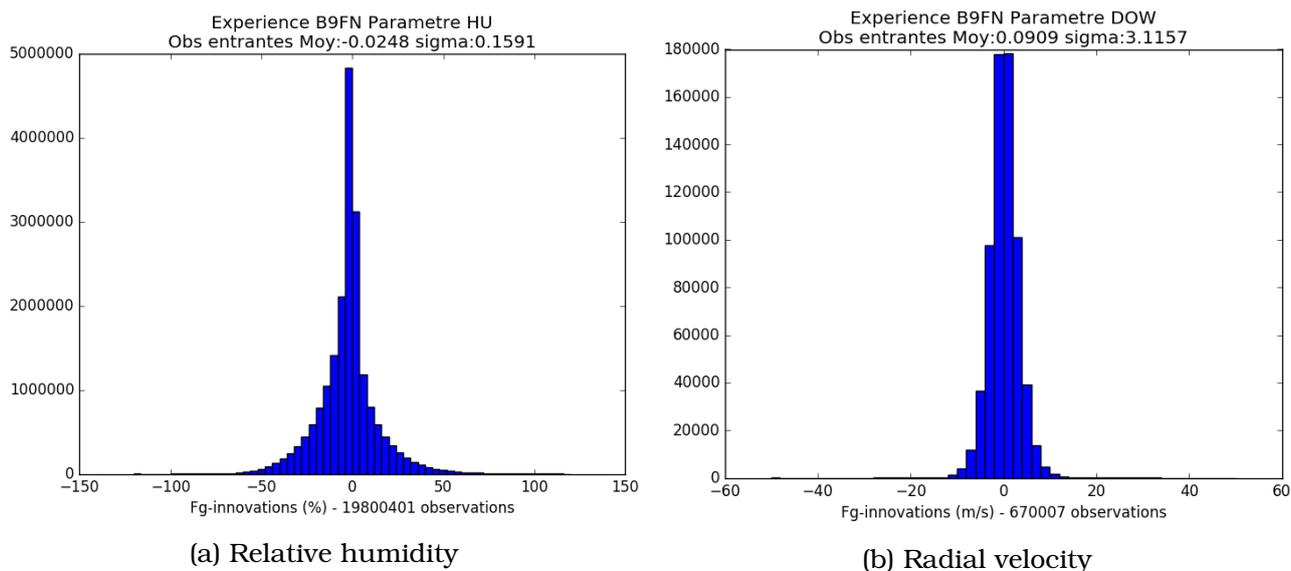


Figure 5: Histograms of innovations of OPERA radars data between 2019/01/24 and 2019/02/02.

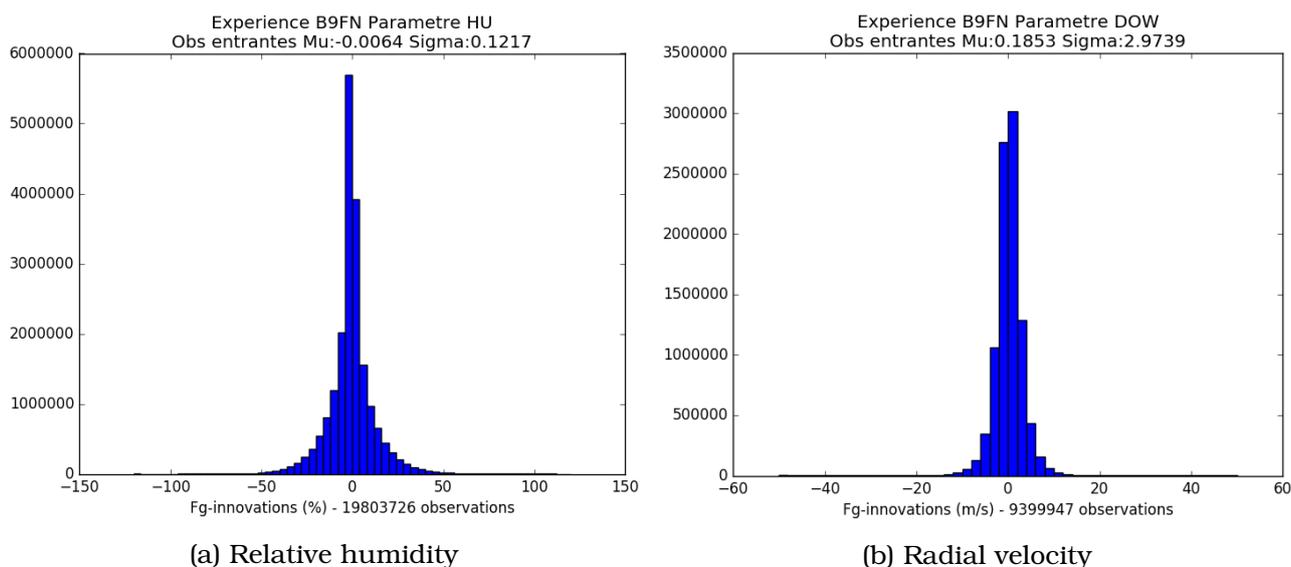


Figure 6: Histograms of innovations of Météo France radars data between 2019/01/24 and 2019/02/02.

3.3 OPERA radar data assimilation: one month experiment

3.3.1 Period of interest

Two AROME-France simulations are compared between 24/01/2019 and 01/03/2019:

- **DBLE**: the operational version with assimilation of 30 Météo France radars (in-house product),
- **B9FY**: the same configuration adding active assimilation of 62 OPERA radars

3.3.2 "Score indicateur"

The "score indicateur" is a mean (on thresholds and neighbourhoods) of Brier Skill Scores against persistence on the parameters accumulated 6 hours rainfall RR6 and averaged wind gust in six hours between the six and twenty-four hours forecast ranges FXI6. For rainfall, thresholds considered are 0.5, 2 and 5 mm; for wind gusts, thresholds considered are 40, 60 and 80 km/h. Scores on two meter temperature T2M and SEVIRI brightness temperature BTP6 are also calculated. Observations are from French ground stations. Figure 7 represents the "score indicateur" for the two experiments over the 31 days period of interest. From this score, the impact of the assimilation of the OPERA radar data is quite neutral.

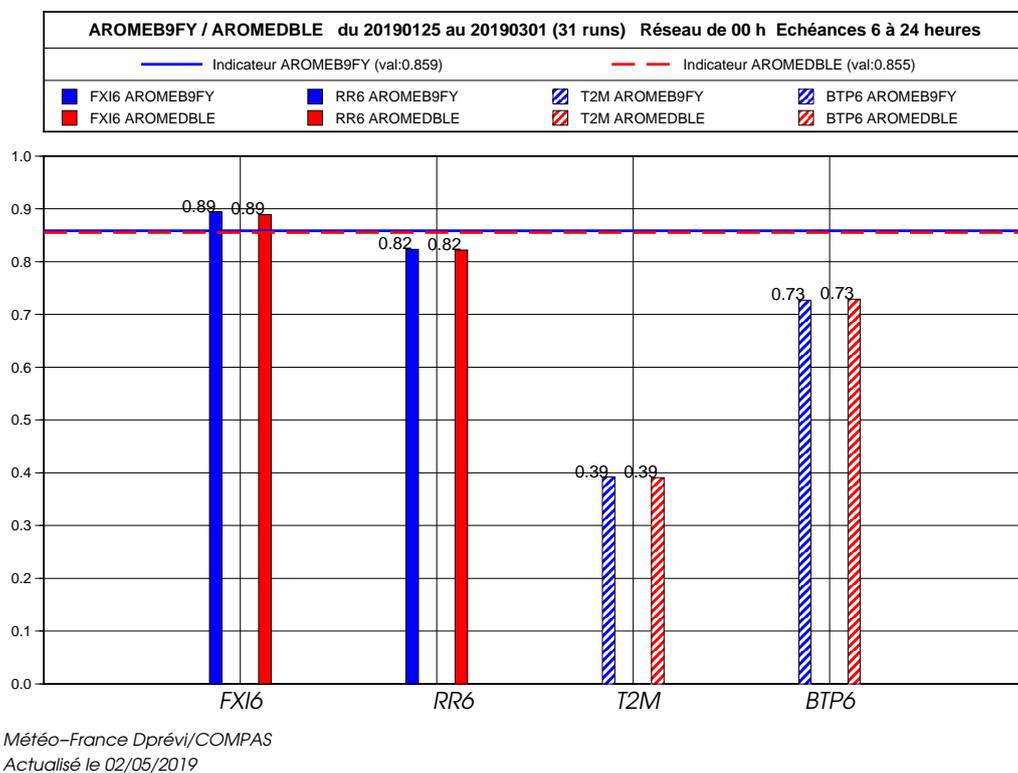


Figure 7: "Score indicateur" compared between the two AROME-France experiments (red: French radars only, blue: French and OPERA radars) between 25/01/2019 and 01/03/2019.

3.3.3 Comparison with ground stations over France

Around 1000 surface stations measure temperature and precipitation and around 700 stations measure wind and humidity over France. Figure 8 represents bias and RMSE for humidity, temperature, wind intensity and precipitation cumulated over 6 hours as a function of forecast time. Scores on temperature, humidity and wind intensity are nearly the same for the two experiments but differences occur on precipitation. Adding OPERA radar data improves the quality on rain accumulations for the first six to twelve hours but a slight degradation is observed at forecast ranges 21 and 24 hours.

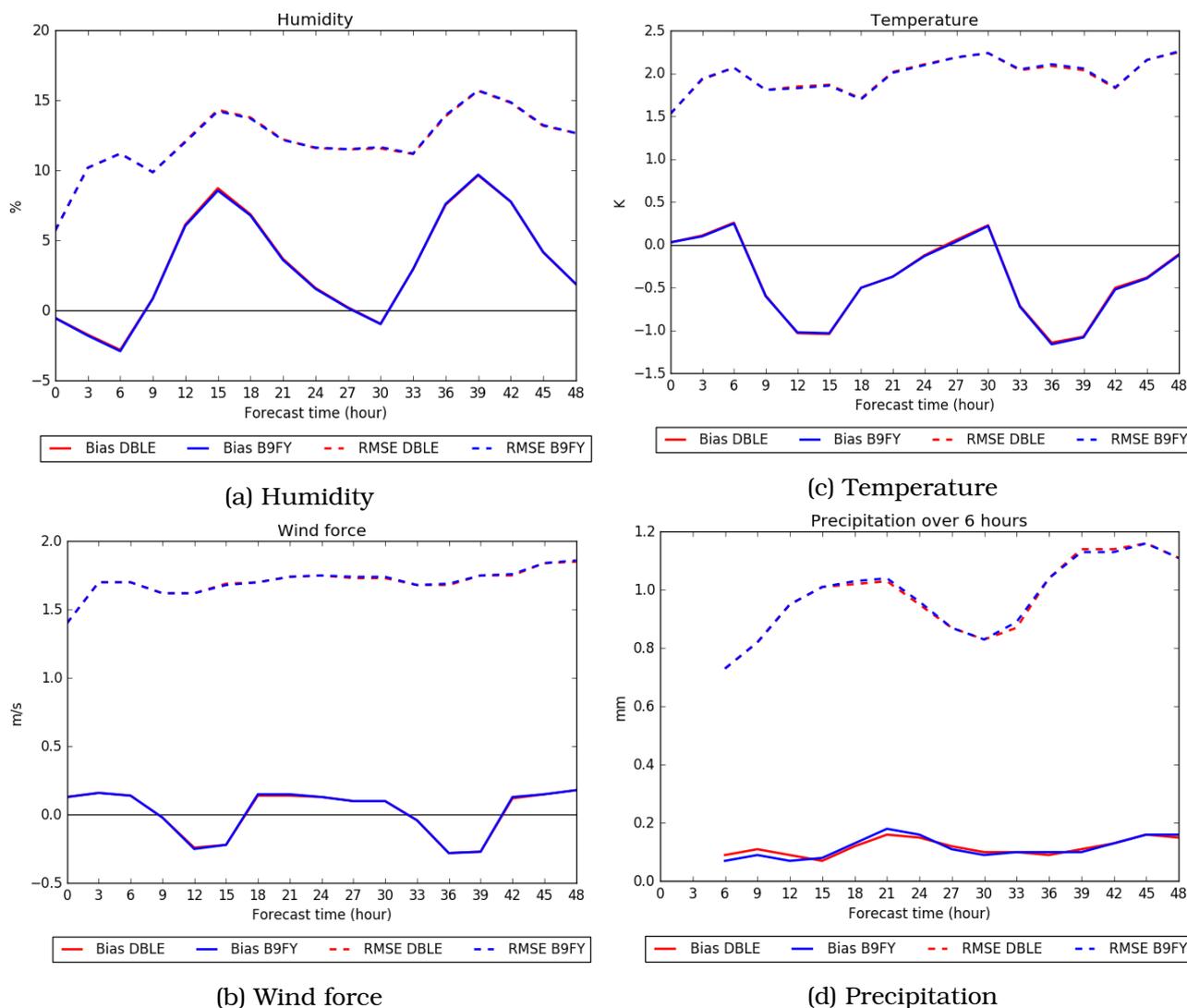


Figure 8: Bias and Root mean square error for ground-station measurements, function of forecast time. Reference in red and experiment with assimilation of OPERA radar data in blue.

3.3.4 Comparison with European ground stations

The number of European surface observations available at Météo France is not as important as Météo France ground stations but it allows to quantify the contribution of OPERA radar data assimilation in AROME-France over Europe outside France. Figure 9 represents ROC curve and ETS for 24 hours precipitation. Four thresholds are considered: 0.5, 2, 5 and 10 mm. The 24 hour accumulation is calculated between forecast time 6 h and 30 h. ROC curve and area under this curve (specified in the legend) show that the impact of OPERA radar data assimilation is rather neutral. ETS indicates a slight improvement for 2

mm threshold and this difference is significant (the blue spot indicates the significance of the difference). Figure 10 represents ROC curve and ETS for 10 m wind gusts observations. 1 h forecast time is considered for this parameter. A slight improvement of wind gusts forecast is observed for strong winds (above 20 m/s) but the differences are not significant. This was expected given the reduced number of radial velocity observations currently used from OPERA.

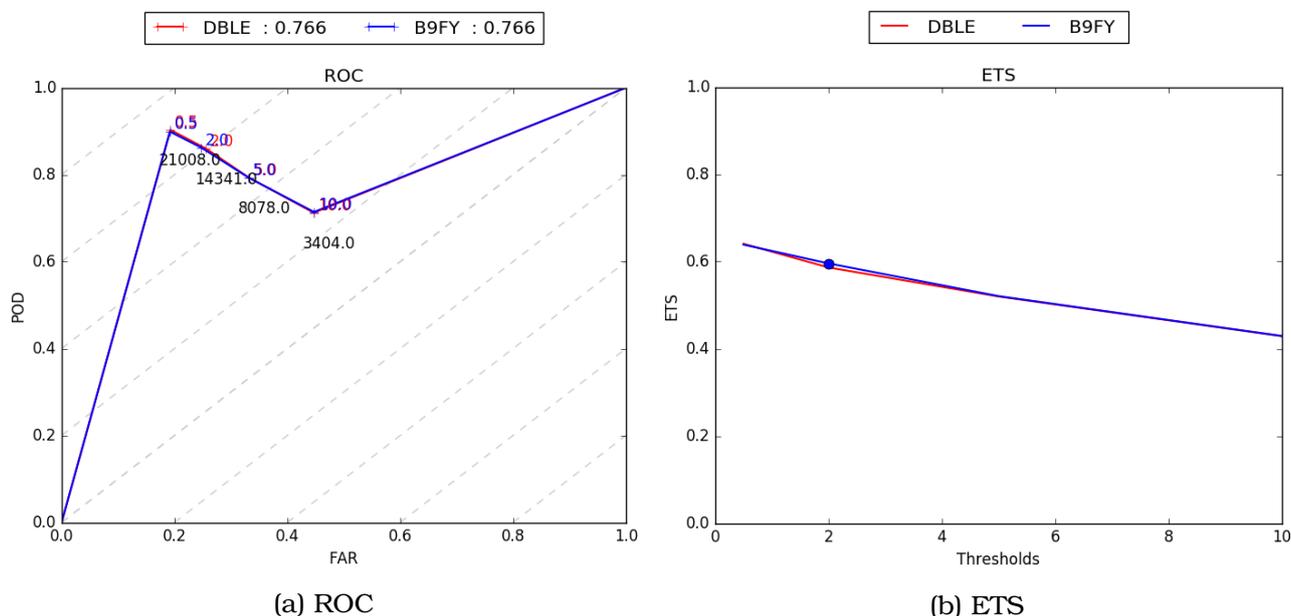


Figure 9: Scores on 24h precipitation for 0.5, 2, 5 and 10 mm thresholds. Forecast time: 6 to 30 hours.

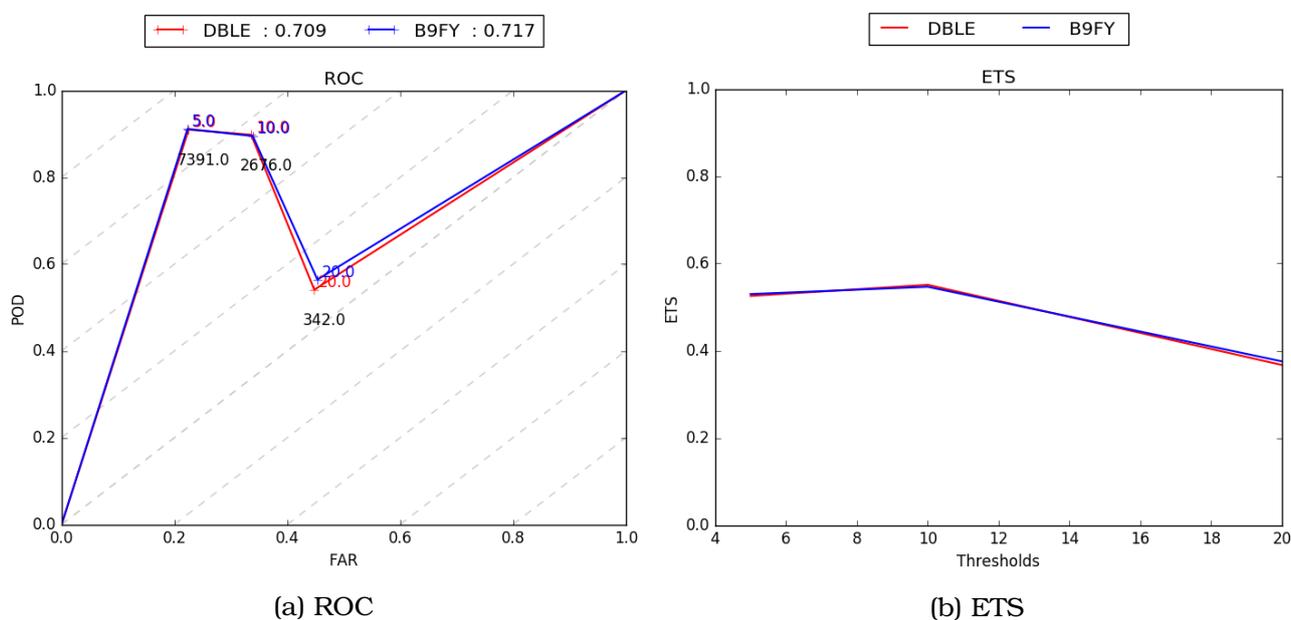


Figure 10: Scores on 10 m wind gusts for 5, 10 and 20 m/s. Forecast time: 1h.

Conclusion

OPERA provides single-site radar data for assimilation purposes through OIFS. 160 radars from 25 countries are available on this ftp server. In AROME-France 3D-Var, reflectivity and radial wind from 30 Météo France ARAMIS radars are already assimilated since 2008, using a specific product. In addition to these data, the goal is to use data from the 62 OPERA radars contributing in AROME-France domain.

The first step in order to correctly use these data is to understand what data are sent and more specifically the quality index produced at Odyssey level. A qualitative study has been undertaken to determine the best threshold to specify in order to suppress non-meteorological data from observations to be assimilated. This study led to a value of 0.7 for the total quality index.

After a monitoring phase showing that the innovations follow Gaussian laws, both for relative humidity and radial wind, an experiment with assimilation of 62 OPERA radars has been performed. The results are quite encouraging even if the impact is rather neutral. One reason is that we have rejected a lot of data from technical reasons (no TH or bad Nyquist velocities).

In order to improve these results, OPERA has to improve the data made available to the NWP community. We keep on asking to have TH PPIs and also have correct metadata in the files. As OPERA 5 has begun, we have to pay attention to possible changes that might happen in the new production streams and we have to be explicit in what data are required for NWP applications. A specific concern is on the quality index. Any change on this parameter could have an impact on assimilation performances and as users, we must ensure that no regression will happen in the current OPERA program.

References

- [Saltikoff et al., 2019] Saltikoff, E., Haase, G., Delobbe, L. Gaussiat, N., Martet, M, Idziorek, D., Leijnse, H., Novák, P., Lukach, M. and Stephan, K. (2019). OPERA the Radar Project. *Atmosphere*, 10(6), 320.
- [Brousseau et al., 2016] Brousseau, P., Seity, Y., Ricard, D., and Léger, J. (2016). Improvement of the forecast of convective activity from the AROME-France system. *Quarterly Journal of the Royal Meteorological Society*, 142.
- [Wattrelot, 2016] Wattrelot, E. (2016). Radar assimilation at higher density in the operational AROME model at 1.3 km horizontal resolution. *Combined Newsletter of the HIRLAM and ALADIN Consortia*, (6):53–63.
- [Wattrelot et al., 2014] Wattrelot, E., Caumont, O., and Mahfouf, J.-F. (2014). Operational implementation of the 1D+3D-Var assimilation method of radar reflectivity data in the AROME model. *Monthly Weather Review*, 142(5):1852–1873.

A Reading ODIM files in BATOR, available from cy43

How Bator does read OPERA Radar files processed by Odyssee (HDF5)

Bator is able to read OPERA radars HDF5 files processed by Odyssee. These files must respect EUMETNET OPERA v2.0, v2.1, or v2.2 information models for implementation with the HDF5 file format. Furthermore, only PVOL and SCAN data file are handled.

1. Used writing rules

- HDF5 keywords are written in **bold**.
- Label used in OPERA radar files are written in *italic*.
- Variables and symbols used in Bator are written using Courier New.

2. Validation of the file and memory allocations

This operation (performed by `PrefetchHdf5()` and `ValidOdin()` subroutines) requires the `param.cfg` and `NAMELIST` files (see documentation concerning these two files for more information). It is composed of the following steps :

- selection of the appropriate template (in `param.cfg` file),
- check that the *Conventions* **attribute** matches any of the allowed values,
- count the number of elevations found in the file, get all *nrays*, *nbins*, *rscaler*, and *rstart* **attributes** in order to allocate the required memory for the `ZENT`, `ZENTSUP`, `ZWAGON` arrays (in `Bator.F90`).

3. Getting required data from file

A *PVOL* file may contain several *dataset*, with different *startdate*, *starttime*, *nrays* **attributes**, which contain one or more required data types (*DBZH*, *TH*, *VRAD*,...) at the same elevation. In this case, we have to choose one (the closest from the analysis date) to get a proper cylinder of observations. This task is one of the aims of the first part of the `odin()` subroutine. The different stages decided at MF in order to select observations are listed below.

a) Required top level attributes.

These required **attributes** are components of *what*, *where*, and *how* top level **groups**. They are stored in the `Radar` structure.

 **If one top level attribute is missing, the data file will be rejected.**
Only NOD identifier is considered and must be defined in source attribute.

b) Other top level attributes

When we read a *SCAN* data file, the OPERA convention allows to have **attributes** which are specific to *dataset* and *data* **groups** at the top level, as supplemental components of *what* and *where* top level **groups**. So, Bator gets these **attributes** (`GetDAttributes()` subroutine) if they exist and stores them in the `Radar%Attrib` structure.

c) Filling the `FullDatasetList` structure

Bator parses the data file getting all **attributes** from *dataset*, *data* and *quality* **groups** to store them in the `FullDatasetList` array.

When parsing ends, all components of `FullDatasetList()%Gdata()%Attrib` and `FullDatasetList()%Gquality()%Attrib` structures are filled.

d) Selection of the most popular *nrays*

- Populates each *nrays* value found with matching *data groups* whose *quantity* takes the value *DBZH, TH, VRAD, or VRADH*.
- Selects the 2 most “populated” *nrays* which must be proportional.
- Keeps the *data groups* matching the selected *nrays* values (and then their elevations). The others are rejected and the corresponding `FullDatasetList() %Gdata() %Attrib` are reinitialized.

e) Selection of the closest elevations to the analysis date

When parsing the `FullDatasetList()GData` array, if several *dataset groups* have the same *elevation* value, only the closest to the analysis date is kept. Bator uses the `SelectedElangles` array to store the result.



The resulting `SelectedElangles` for a given elevation value will be a mix of the different quantities found in the datasets groups which match this elevation.

f) Elevations Sort, getting data, and thinning along rays

Sorts selected elevations in the ascending order, gets the data of each *quantity* (and associated flags) and thins them along the ray according to the required resolution. The `Radar%FinalElev` array is used to store the result. This array has to be used in the second part of the `odim()` subroutine.