

1) Most common filters/processing for non-satellite data

Here is the numbered list of the controls made for almost all data types. For each data type, the numbered control has to apply when its cell is checked.

1. Rejecting observations when out of sphere.
2. Rejecting observations when altitude is missing.
3. Rejecting observations when date is outside the assimilation temporal window.
4. Rejecting observations when identification is missing.
5. Formatting identification.
6. Rejecting whole observation when no valid parameter is found.
7. DD/FF to U/V conversion.

sensor	Filter/processing									
	1	2	3	4	5	6	7	8	9	10
Gpssol	X	X	X	X	X	X ¹				
Radar	X		X			X ²				
Synop	X		X			X	X			
Acar	X	X	X			X	X			
Airep/Amdar	X		X			X	X			
buoy	X		X			X	X			
Paob/paobvent	X		X			X ³				
europofil	X		X			X				
pilot	X		X			X ⁴				
profiler	X	X	X			X				
radomeh	X		X			X				
synor	X		X			X				
temp	X					X				
Bathy/tesac	X		X			X				

2) Gpssol (bufr - only Bator)

Need a whitelist is needed (for each model) with the following input format :

```

statid      lat      lon      alt  dt      bias       $\sigma_{mon}$  (mm)   $\sigma_{obs}$  (m.)
ACNSIGE_   41.70   -6.35   815.  15.    0.005560  10.02    0.00802
    
```

This file is generated from monitoring data (obs-guess, monthly).

- Each observation is first flagged as “passive” for monitoring.
- “Active” observations are not created in the following cases :
 - when latitude, longitude, period, altitude do not match those specified in the whitelist,

Météo-France : filters & processing used in OULAN/BATOR (based on CY38t1_op1)

- when minutes of observation date do not match 00,15,30,or 45 for METO, SGN_, SGN1 and ASI_ identifiers, or match 59 for the others identifiers,
- when the observation value ('atmospheric path delay') $\notin [1.8, 3[$
- For Arpege, we average the 'atmospheric path delay' of 'whitelisted' statids for each timeslot, and allocate date and time values of the timeslot's center. So, a new active observation is created unless it already exists as passive one with the same temporal data ; in such case, we activate it and update its data.
- For Aladin/Arome, we activate the 'whitelisted' statid observation for which the date is the closest to central date of the timeslot.

3) Radar (bufr - only bator)

- Data extraction (one BUFR file per elevation) must be ordered by increasing elevation for each radar station.
- An elevation is rejected when $\notin [0, 80]$ or when elevation time is missing.
- We consider two coordinate systems : polar system and Conical Cartesian system.
- An observation is defined as a vertical profil. This consists of the set of assumed collocated data from existing elevations.
- Quality flag is used to keep or reject each parameter (radial velocity and reflectivity).
- Reflectivity values 'at noise level' are set to a MDZ (minimum detectable reflectivity) threshold. All reflectivities are corrected from attenuation caused by precipitations, when possible (especially for dual polarization radars).
- Three successive filters are applied to radial velocities :
 - we count the pixels (for a given elevation) whose deviation from the average of the neighboring pixels is greater than a prescribed threshold. The whole elevation is rejected when the number of pixels is too large (threshold en %).
 - When the elevation is not rejected, we applied a median filter.
 - At last, a pixel takes the average value of neighboring pixels when the deviation between its observed value and the computed average is greater than a threshold.

4) Synop/synop-ship (obsoul)

- Anemometer height value is 10m for land synop.
For 'ship' reports the real value is taken when available in the message, otherwise a default value of 24m is considered.
- Relative humidity and specific humidity are computed (from T2m & Td).

5) Acar (obsoul)

- Altitude can be in meters or Pa.

6) Airep (obsoul)

- Altitude in meters.

Météo-France : filters & processing used in OULAN/BATOR (based on CY38t1_op1)

- Relative humidity is computed from T2m & Td (when not missing).

7) Buoy (obsoul)

- Relative humidity is computed from T2m & Td (when not missing).

8) Paobvent (obsoul)

- A preprocessing is made before bator :
- its purpose is to create a three dimensional vortex of cyclonic winds according the analytical model of Holland (1980). Two circles of data at 100 and 200 km from the center of the cyclone (and an additional one at 10km+radius of max wind, if radius<=60km), at 10m , 850, 700 and 500 hPa are created.
8 wind vectors are produced by circle and level.
Vortex 10m winds are computed from the center PMER, 10m max wind and its radius. An attenuation factor (depending on height) is used for the upper level winds. A convergence of 12° is applied to the 10m wind to take into account the surface friction.
The motion vector of the cyclone (during the last six hours), is added to the computed wind vectors to take into account the cyclone's asymmetries.

9) Radomeh (obsoul)

- Relative humidity and specific humidity are computed (from T2m & Td).

10) Temp

- Relative humidity and specific humidity are computed (from T2m & Td).

- (1) When time significance code ≤ 23 (monitoring period), elevation $\leq 90^\circ$, or time period $\notin [5, 60]$ (minutes).
- (2) When no suitable BUFR files (number of subsets ≤ 1 , or 450, or 512).
- (3) When producer code $\notin \{25, 34, 52, 65, 71\}$, or when all others data are missing.
- (4) When pressure and geopotential are missing for one level, or when no wind data found.