BUFR, NETCDF & HDF5
DATA PREPROCESSING
FOR ARPEGE/ALADIN/AROME.

PARAM.CFG FILE (BATOR)

VERSION ANGLAISE / ENGLISH VERSION

V. 1.0.0

CONTENTS

1 Fist of all

1.1 Used writing rules

- File names, jobs and programs are written in **bold** font.
- Block and key labels used in **param.cfg** file are written in *italic*.
- Examples are written using the `Courier New` font.
- In examples, optional entries are inserted within square brackets.
- Likewise, in examples, items written using the *`Courier New italic`* font must be replaced by their values.

2 Introduction

The purpose of this guide is to explain how to write a **param.cfg** file, used as an input file for **Bator** from CY43.
Knowledge of BUFR, NETCDF and HDF5 files formats is required.

3 History

Version 1.0.0 (07/07/2017) :
– first version of this document in French, valid with CY43T2.
– this release translated into English (18/05/2018).

4 Aim of the param.cfg file

This file consists of blocks containing the required records allowing **bator** to recognize and process (or reject !) an input data file (BUFR, NETCDF or HDF5 format).
This implies that a particular block and its records are related to the format of a particular file and the data-type it contains.
Furthermore, some records may be added to be used in decoding subroutines of **bator** (following the programmer's own choices !).

5 Common writing rules

- No blank is allowed before a record or a block label.
- Length of records can be up to 255 characters.
- Items *indice*, *posbufr*, *valeur*, *number* and *FXY* are integers.
- Items *chaîne* is a string and must be surrounded with ‘’ .
- A comment line must begin with #.
- A comment cannot follow a record on the same line.
- Items of a record are separated by one or more blank character.
- The file is case sensitive.

6 For BUFR files

6.1 Description block syntax

```
BUFR label
nb1 nb2 nb3 nb4
codage      1     +/-FXY1 [information]
...
[codage  indicen +/-FXYn [information]]
{control    1     valeur [information]]
...
[control indicen valeur [information]]
[offset     1     valeur [information]]
...
[offset  indicen valeur [information]]
[values  posbufr FXY [information]]
...
[values  posbufr FXY [information]]
/BUFR label
```

6.2 BUFR label – /BUFR label (or BEGIN label – END label)

Respectively open and close a description block for BUFR file which data-type is given by *label*. *label* is a string up to 16 characters and must match one of the listed sensors in the ExpandBufrFile() subroutine (see bator_decodbufr_mod.F90).
For backward compatibility with CY42, previous writings, *BEGIN label* and *END label* are also allowed.

It is allowed to create several description blocks for a particular data-type. In this case, the list of « codage » keys must be different from one block to another one.

6.3 nb_1 nb_2 nb_3 nb_4

These values give number of records *codage*, *control*, *offset*, *values* contained in the block description :

- *nb1* for number of records for *codage*. Must be > 0.
- *nb2* for number of records for *control*. Must be ≥ 0.
- *nb3* for number of records for *offset*. Must be ≥ 0.
- *nb4* for number of records for *values*. Must be ≥ 0.

6.4 encoding (key “codage”)

Bator uses this key to recognize a BUFR file. For a particular file, the list of *codage* keys must strictly match the list (in number and order) of unexpanded descriptors inside the file (section 3 of a BUFR file). So, there is at least one *codage* key.
In some cases, BUFR files which contain the same data-type may have one or several different complementary unexpanded descriptors before and/or after the main sequence. To minimize number of blocks contained in **param.cfg** file, you can list these unexpanded descriptors preceding them with a minus character. If applied, these unexpanded descriptors will be ignored. See example page 13.

6.5 control

Optional key used to give an additional information which is missing in the BUFR file (number of channels, delayed replication factor position when more than one exists…). This facility may be used to enable dedicated decoding software to be more flexible.

6.6 offset

The value defined by this optional key gives the number of B descriptors to jump when an iteration is done in the decoding subroutine.

6.7 values

The values associated with this key give the position and the FXY of a descriptor (section 4 of a BUFR file) which value must be retrieved in the decoding subroutine.
Although its use is optional, it is strongly recommended.
By convention, if *posbuf* value is set to -1 for a given descriptor, its position must be calculated in the decoding subroutine,

 *A descriptor (FXY) must be unique. Nevertheless, if it is listed several time, only the last one will be considered.*

6.8 Information field

This optional field, unused by **bator**, can be useful for the programmer. It gives information concerning the current record (WMO label…).

7 For NETCDF(4) Files

This file format is currently used for Lannion SEVIRI data preprocessing. The syntax of this description block can evolve when other data-type preprocessing will be planned. See example page 16.

7.1 Description block syntax

```
NETCDF label
nb1 nb2 nb3 nb4
dimension  'chaîne1'
...
[dimension 'chaînen']
{genattrib 'chaîne1']
...
[genattrib 'chaînen']
variable   'chaîne1'
...
[variable  'chaînen']
[channel   'chaîne1']
...
[channel   'chaînen']
/NETCDF label
```

7.2 NETCDF label – /NETCDF label

Respectively open and close a description block for NETCDF4 file which data-type is given by *label*. *label* is a string up to 16 characters and must match one of the listed sensors in PrefetchNetcdf() subroutine (see bator_decodnetcdf_mod.F90).

a NETCDF description block must list all the first level items contained in the data file.

7.3 nb_1 nb_2 nb_3 nb_4

These values give the number of records of *dimension*, *genattrib*, *variable*, *channel* expected in the description block :

- *nb1* for number of records for *dimension*. Must be > 0.
- *nb2* for number of records for *genattrib*. Must be ≥ 0.
- *nb3* for number of records for *variable*. Must be > 0.
- *nb4* for number of records for *channel*. Must be ≥ 0.

7.4 dimension

The value of this key, *chaîne*, matches a dimension label used in the NETCDF data file.
There must be as many records using this key as there are dimension labels in the NETCDF data file.

7.5 genattrib

The value of this key, *chaîne*, matches a global attribute label used in the NETCDF data file.

7.6 variable

The value of this key, *chaîne*, matches a variable label used in the NETCDF file. It is used to get meta-data like latitude, longitude and so on.

7.7 channel

Like *variable* key, the value of this key, *chaîne*, matches a variable label used in the NETCDF file. Its particularity is to concern data for a particular channel, instead of metadata.

8 For HDF5 files

8.1 Description block syntax

Syntaxe :

```
HDF5 label
nb1 nb2 nb3
[groupe    number  'chaîne1']
...
[groupe    number  'chaînen']
{genattrib      0  'chaîne1']
...
[genattrib      0  'chaînen']
[dataset   number  'chaîne1']
...
[dataset   number  'chaînen']
/HDF5 label
```

8.2 HDF5 label – /HDF5 label

Respectively open and close a description block for HDF5 file which data-type is given by *label*. *label* is a string up to 16 characters and must match one of the listed sensors in PrefetchHdf5() subroutine (see bator_decodhdf5_mod.F90).

a HDF5 description block must list all the first level items contained in the data file.

8.3 nb_1 nb_2 nb_3

These values give number of records *groupe*, *genattrib*, *dataset* contained in the block description :

- *nb1* for number of records for *groupe*. Must be ≥ 0.
- *nb2* for number of records for *genattrib*. Must be ≥ 0.
- *nb3* for number of records for *dataset*. Must be ≥ 0.

8.4 groupe

This key needs two parameters :

- *number* is a unique identifier (integer) affected to the group and chosen by the programmer.
- *chaîne* is the label (or the constant part of the label) of the group in the HDF5 file.

If « chaîne » is the constant part of a group label, which must be at the beginning of this label, « number » must be negative. See example page 18.

8.5 Genattrib

The value of this key, *chaîne*, matches a top level attribute label used in the HDF5 file.
Constant 0 is mandatory.

8.6 dataset

This key needs two parameters :

- *number* is a unique identifier (integer) affected to the dataset and chosen by the programmer.
- *chaîne* is the label (or the constant part of the label) of the group in the HDF5 file.

If « chaîne » is the constant part of the dataset label, which must be at the beginning of this label, « number » must be negative. See example page 18.

ANNEXES

ANNEXE – 1 : PARAM.CFG FILES

1 For some BUFR files

```
BUFR synop1
3 2 2 27
codage     1  208035
codage     2  307080
codage     3  208000
control    1      42 second delayed replication factor's position
control    2       2 iteration's number for RR
offset     1       4 nb of replicated B descriptors
offset     2       5 nb of replicated B descriptors
values    -1  007063 Depth below sea/water surface
values    -1  022043 Sea/water temperature
values    -1  001101 State identifier
values     1  001001 WMO block number
values     2  001002 WMO station number
values     3  001015
values     4  002001 Type of station
values     5  004001 Year
values    10  005001 Latitude
values    11  006001 Longitude
values    12  007030 Height of station ground above mean sea level
values    13  007031 Height of barometer above mean sea level
values    14  010004 Pressure
values    15  010051 Pressure reduced to mean sea level
values    19  007004 Pressure (standard level)
values    20  010009 Geopotentiel height of the standard level
values    22  012101 Temperature/dry-bulb temperature
values    23  012103 Dew_point temperature
values    28  013023 Total precipitation past 24 hours
values    30  020010 Cloud cover (total)
values    37  031001 Delayed replication factor (the first in file)
values    61  013013 Total snow depth
values    72  004024 Time period in hours
values    73  013011 Total precipitation/total water equivalent of snow
values    86  004025 Time period (-10, or number of minutes after change)
values    87  011001 Wind direction
values    88  011002 Wind speed
/BUFR synop1

BUFR synop1
1 2 2 27
codage     1  307080
control    1      42 second delayed replication factor's position
control    2       2 iteration's number for RR
offset     1       4 nb of replicated B descriptors
offset     2       5 nb of replicated B descriptors
values    -1  007063 Depth below sea/water surface
values    -1  022043 Sea/water temperature
values    -1  001101 State identifier
values     1  001001 WMO block number
values     2  001002 WMO station number
values     3  001015 Station or site name
values     4  002001 Type of station
values     5  004001 Year
values    10  005001 Latitude
values    11  006001 Longitude
values    12  007030 Height of station ground above mean sea level
values    13  007031 Height of barometer above mean sea level
values    14  010004 Pressure
values    15  010051 Pressure reduced to mean sea level
```

```
values    19  007004 Pressure (standard level)
values    20  010009 Geopotentiel height of the standard level
values    22  012101 Temperature/dry-bulb temperature
values    23  012103 Dew_point temperature
values    28  013023 Total precipitation past 24 hours
values    30  020010 Cloud cover (total)
values    37  031001 Delayed replication factor (the first in file)
values    61  013013 Total snow depth
values    72  004024 Time period in hours
values    73  013011 Total precipitation/total water equivalent of snow
values    86  004025 Time period (-10, or number of minutes after change)
values    87  011001 Wind direction
values    88  011002 Wind speed
/BUFR synop1

BUFR hirs
1 1 0 15
codage     1  310008
control    1      19  nb de canaux
values     3  001034  subcentre
values     7  001007  Satellite identifier
values    12  005043  Field of view number
values    11  005041  Scan line number
values    22  005001  Latitude
values    23  006001  Longitude
values    16  004001  Year
values    24  007001  Height of station
values     8  002048  Satellite sensor type
values    25  007024  Satellite zenith angle
values    26  005021  Bearing or azimuth
values    27  007025  Solar zenith angle
values    28  005022  Solar azimuth
values    38  002150  Tovs Channel number
values    43  012063  Brightness Temperature
/BUFR hirs

BUFR amdaromm
5 0 3 16
codage     1  311010
codage     2  -004015
codage     3  -102004
codage     4  -004032
codage     5  -011037
offset     1  2
offset     2  0
offset     3  0
values     3  001006  Aircraft flight number
values     1  001008  Aircraft registration number
values    21  005001  Latitude
values    23  006001  Longitude
values     9  004001  Year
values    25  007010  Flight level
values    -1  008004  Phase of flight
values    29  008009  Detailed phase of flight
values    31  011001  Wind direction
values    33  011002  Wind speed
values    47  012101  Temperature/dry-bulb temperature
values    51  013002  Mixing ratio
values    56  012103  Dew-point temperature
values    -1  002061
values    54  031000
values    -1  031001
/BUFR amdaromm

BUFR amdaromm
1 0 3 16
codage     1  311010
offset     1  2
```

```
offset     2  0
offset     3  0
values     3  001006  Aircraft flight number
values     1  001008  Aircraft registration number
values    21  005001  Latitude
values    23  006001  Longitude
values     9  004001  Year
values    25  007010  Flight level
values    -1  008004  Phase of flight
values    29  008009  Detailed phase of flight
values    31  011001  Wind direction
values    33  011002  Wind speed
values    47  012101  Temperature/dry-bulb temperature
values    51  013002  Mixing ratio
values    56  012103  Dew-point temperature
values    -1  002061
values    54  031000
values    -1  031001
/BUFR amdaromm
```

2 For SEVIRI NETCDF file

```
NETCDF seviri
2 36 17 8
dimension  'nx'
dimension  'ny'
genattrib  'satid'
genattrib  'ncml_version'
genattrib  'nwc_saf_algorith_version'
genattrib  'title'
genattrib  'comment'
genattrib  'references'
genattrib  'created_by'
genattrib  'institution'
genattrib  'source'
genattrib  'Metadata_Conventions'
genattrib  'summary'
genattrib  'hrit_converted_by'
genattrib  'keywords'
genattrib  'keywords_vocabulary'
genattrib  'naming_authority'
genattrib  'cdm_data_type'
genattrib  'creator_name'
genattrib  'creator_url'
genattrib  'creator_email'
genattrib  'geospatial_vertical_min'
genattrib  'geospatial_vertical_max'
genattrib  'time_coverage_start'
genattrib  'time_coverage_end'
genattrib  'time_coverage_duration'
genattrib  'time_coverage_resolution'
genattrib  'licence'
genattrib  'Conventions'
genattrib  'Area_of_acquisition'
genattrib  'Scanning_direction'
genattrib  'history'
genattrib  'date_created'
genattrib  'id'
genattrib  'geospatial_lat_min'
genattrib  'geospatial_lat_max'
genattrib  'geospatial_lon_min'
genattrib  'geospatial_lon_max'
variable   'lat'
variable   'lon'
variable   'sat_azi_ang'
variable   'sat_zen_ang'
variable   'CTP'
variable   'CT'
variable   'CTP_QUALITY'
variable   'CT_QUALITY'
variable   'time'
variable   'dtime'
variable   'commentaires'
variable   'satellite'
variable   'geos'
variable   'ImageNavigation'
variable   'GeosCoordinateSystem'
variable   'Y'
variable   'X'
channel    'IR_039'
channel    'WV_062'
channel    'WV_073'
channel    'IR_087'
channel    'IR_097'
channel    'IR_108'
channel    'IR_120'
channel    'IR_134'
```

```
/NETCDF seviri
```

3 For HDF5 files

```
HDF5 mtvza
0 1 31
groupe    1 'HDF4_DIMGROUP'
dataset  -1 'm_m2_01_10.6V'
dataset  -2 'm_m2_02_10.6H'
dataset  -3 'm_m2_03_18.7V'
dataset  -4 'm_m2_04_18.7H'
dataset  -5 'm_m2_05_23.8V'
dataset  -6 'm_m2_06_23.8H'
dataset  -7 'm_m2_07_36.7V'
dataset  -8 'm_m2_08_36.7H'
dataset  -9 'm_m2_09_91.65V'
dataset -10 'm_m2_11_52_80V'
dataset -11 'm_m2_12_53_30V'
dataset -12 'm_m2_13_53_80V'
dataset -13 'm_m2_14_54_64V'
dataset -14 'm_m2_15_55_63V'
dataset -15 'm_m2_16_57_0.32_0.1H'
dataset -16 'm_m2_17_57_0.32_0.05H'
dataset -17 'm_m2_18_57_0.32_0.025H'
dataset -18 'm_m2_19_57_0.32_0.01H'
dataset -19 'm_m2_20_57_0.32_0.005H'
dataset -20 'm_m2_21_183_7.0V'
dataset -21 'm_m2_22_183_1.4V'
dataset -22 'm_m2_23_183_3.0V'
dataset -23 'm_m2_26_31.5V'
dataset -24 'm_m2_27_31.5H'
dataset -25 'm_m2_Julian Day'
dataset -26 'm_m2_Lattitude'
dataset -27 'm_m2_Longitude'
dataset -28 'm_m2_SunAzimuth'
dataset -29 'm_m2_SunZenith'
dataset -30 'm_m2_Time of day'
dataset -31 'm_m2_Surface'
/HDF5 mtvza

HDF5 odim
1 4 0
genattrib  0 'Conventions'
groupe     1 'what'
groupe     2 'where'
groupe     3 'how'
groupe    -4 'dataset'
/HDF5 odim
```