

How to split BUFR subsets

A BUFR message can be composed of several subsets. The number of subsets in a BUFR message is given by the key "numberOfSubsets" which is printed by the bufr_ls tool:

```
> bufr_ls synop_multi_subset.bufr
centre    masterTablesVersionNumber  localTablesVersionNumber
numberOfSubsets
enmi      14                      0                          12
```

Starting from release version 0.13.0 a given number of subsets can be extracted and saved in a new message. This is based on the use of the following keys:

- extractSubset
- extractSubsetIntervalStart
- extractSubsetIntervalEnd
- doExtractSubsets

To explain the use of these keys the following example using the bufr_filter is commented, but we highlight the fact that the keys can be set using the Fortran and python interfaces as well and no special functions have been implemented in the bufr_filter for this feature.

The following file "instructions.filter" with bufr_filter instructions can be used to perform the extraction of several subsets from a BUFR message.

instructions.filter

```
set unpack=1;

set extractSubset=4;
set doExtractSubsets=1;
write;

set extractSubsetIntervalStart=3;
set extractSubsetIntervalEnd=8;
set doExtractSubsets=1;
write;
```

First of all the data section has to be decoded with the instruction at line 1. Setting the key extractSubset will express the subset that is going to be extracted. In this example in line 3 we ask ecCodes to extract the 4th subset from the input file. The extraction is executed when the key doExtractSubset is set to 1 and the message containing only the 4th subset is written to the output file with the write instruction in line 5.

It is possible to extract all the subsets in an interval starting with the key extractSubsetIntervalStart and ending with the key extractSubsetIntervalEnd. In the example at lines 7 to 8 the request to extract all the subsets between the 3rd and the 8th (including the 3rd and the 8th) is expressed setting the two extractSubsetInterval keys. The extraction is performed when the key doExtractSubset is set to 1 and the resulting message containing the 6 subsets from the 3rd to the 8th from the original message are written to the output file.

This example can be used on any multi-subset input file in compressed and uncompressed form as follows:

```
> bufr_filter -o out.bufr instructions.filter in.bufr
```

There is also the key "extractSubsetList" which allows you to specify the subsets of interest in a list e.g.

```
# Extract the first and the third subsets
set extractSubsetList={1, 3};
set doExtractSubsets=1;
write;
```

