QC Check-list for CORDEX

H.-D. Hollweg

DKRZ, Hamburg

Date: 2013-05-30

Introduction

The QC checks for compliance of meta-data and tests for breaks in time value sequences as well as suspected data. This paper discribes how handling of annotations is accomplished.

In fact, testing and issuing annotations are different processes. Testing is done always by the program; the user has no influence (but by coding). However, handling of annotations is controlled from case to case by inspecting the check-list table *QC-0.4/tables/CORDEX_check-list.conf*; default values are always stored in the folder *QC-0.4/tables/SVN_defaults*.

The table specfies a grade of severeness to each type of incident in the data file, whether an annotation is to be discarded, a notification should take place even by email or whether the directives are valid for all variables, a group, or adjusted to a single one. The connection between a failed test in the QC program and the corresponding annotation is given by a unique flag. Also the interface is common to all checks. The proceeding is:

1. a test failed,

2. determination of whether to discard or issue an annotation,

3. detailed information about the test may be passed to the annotation procedure, (if nothing is passed, then the descriptive text of the table is applied),

4. annotations are delivered as well as decisions whether to stop the program immediately, or after completion of meta-data checks, or stipulate prevention of further processing of subsequent files of the same kind, etc. .

Table Description

```
Syntax:  text & [{]token-list[}][, {token-list}, ...]
```

Each deteced incident is handled by a directive on its own. The ampersand & is required for separation. Directives are specified by a comma-separated list of tokens (surrounding spaces are ignored), which in multiple cases have to be given in braces.

{}

   pairing braces indicate a directive. They may be omitted when there is only a single one.

#

   Any text following the # char is a comment.

Valild tokens:

   L1, L2, L3, D, EM, ST, PT, *flag*, *variable-acronym*, V=*value*, R=*num|range*

level: (grade of severeness)

   L1 Notify incidents. Continue unrestricted.
   L2 Continue checking to the end of the file. Then, lock processing of the current variable.
   L3 Stop after completion of meta-data checks. Lock processing of the current variable.
   L4 Stop immediately and shutdown the QC session (no locks).

table:

   Standard output table (**ST**), project table (**PT**).

flag:

   Two types of flags are defined: a leading integer value *num* for meta-data checks and **R***num* for incidents in data. Grouping of some flags is indicated by a post-fix, e.g. *40_1*. Must match the flag in the QC.cpp class.

EM

   Notification by email (**EM**). Note that this will happen only if the configuration option *EMAIL_TO=addresses-list* is enabled.

D

   Discard annotation.

variable:

> Comma-separated acronyms of variables; directive is only applied to the specified variable(s). Omission: all variables.

V=value

> Constraining value, e.g. {flag,D,V=0,var} discards a test for variable *var* only if value=0.

R=num[-num]

> Record(s), i.e. data at a time value, may be excluded from checking. Single records as well as ranges, e.g. 1-12, may be provided.

Default: {ST,L1}, {PT,L2,EM}, {L3,EM}

There are configuration options, which affect issuing of annotations (see *QC-0.4/doc/qc-config.pdf* for details) by overriding the table entries. Note: same rules for braces.

NOTE=*directive*: Supersede an entry of the check-list.

NOTE_ALWAYS=*directive*: Apply a unique annotation handling to all entries of the check-list.

NOTE_LEVEL_LIMIT=*level*: Cap annotation level of every entry of the check-list.

Table Entries

The directives given below are provided for CORDEX as default.

Time/Data (tested at each time value)**:**

Negative time step	R1, L2, EM
Missing time step	R2, L2, EM
Identical time step	R4, L2, EM
Negative/zero time bounds range	R8, L2, EM
Overlapping time bounds ranges	R16, L2, EM
Gap between time bounds ranges	R32, L2, EM
Found a record entirely with filling value	R100, L1

Found a record entirely with constant value	{R200, L1}, {R200, D, V=0, clivi, mrfso, rsus, rsds, snc, snd, snm, snw, sund}
Failed outlier test: global minimum	R400, L1
Failed outlier test: global maximum	R800, L1
Undefined standard deviation	R1600, D
Suspecting a replicated record	{R3200, L1}, {R3200, D, clivi, mrfso, rsus, rsds, rsdt, snc, snd, snm, snw, sund}
Infinite (Inf) or not-a-number (NaN)	R6400, L2, EM

Directory and Filename Structure:

Directory structure does not match global attributes.	10, L2, EM
Filename not compliant with CORDEX encoding.	15_1, L2, EM
Filename does not match global attributes.	15_2, L2, EM
Variable <name> in filename does not match any variable in the file.	15_3, L2, EM
Wrong separator in the period in the filename.	16_1, L2, EM
Period in filename (1st date): misaligned to time values.	16_2, L2, EM
Period in filename (2nd date): misaligned to time values.	16_3, L2, EM
Period in filename: invalid range.	16_4, L2, EM
Period in iflename: incorrectly formatted.	16_5, L2, EM
Period in filename: incorrectly rounded.	16_6, L2, EM
Period in filename: incorrectly aligned for a season.	16_7, L2, EM
Time ambiguity in the ensemble of files.	16_8, L2
Unspecific netCDF or time failure.	16_9, L2

Required attributes:

Auxiliary <name>: missing required attribute <name>.	20_1, L2, EM

Auxiliary <name>:attribute <name>: missing required attribute value=<value>.	20_2, L2, EM
Auxiliary <name>:attribute <name>: unmatched required attribute value=<value>.	20_3, L2, EM
time: units: missing attribute.	21_1, L2, EM
time: units: wrong or missing key-word.	21_2, L2, EM
time: units: missing reference date.	21_3, L2, EM
time: units: wrong reference date.	21_4, L2, EM
time: calendar: missing attribute	21_5, L2, EM
time: time_bnds: variable declared, but not found.	21_6, L2, EM
Missing required global attribute: <name>.	22_1, L2, EM
Global attribute <name>: missing required value=<value>.	22_2, L2, EM
Global attribute <name>: unmatched required value=<value> .	22_3, L2, EM
Auxiliary <plev> for cloud amounts: missing attribute <plev_bnds>.	23, L2, EM

Variable (target):

Variable <name> not found in the standard table.	30, L1
Variable <name> not CF compliant.	31, L1
Variable <name> has wrong data type.	32_1, L2, EM
<variable>: standard name conflict.	32_2, L2, EM
<variable>: long name conflict.	32_3, L2, EM
<variable> has no units attribute.	32_4, L2, EM
<variable> has wrong units attribute.	32_5, L2, EM
<variable>: cell-method conflict.	32_6, L2, EM
Variable <name> matches only for case-insensitivity.	32_7, L2, EM
<variable>: conflict for attribute positive.	32_8, D
Name of a variable has changed for sub-temporal file.	33, L2, EM
<variable>: _FillValue/missing_value does not match required value=1.E20	34, L2, EM
Variable <name>: units not CF compliant.	35, L1
Variable <name>:time: not CF compliant.	35, L1

Variable <name>:time:units: different or missing period.	36_1, L2
Variable <name>:time:units: different or missing reference date.	36_2, L2

Dimensions (only project table)**:**

Conflict for <variable>, <dimension>: output name.	40_1, L2, EM
Conflict for <variable>, <dimension>: standard name.	40_2, L2, EM
Conflict for <variable>, <dimension>: long name.	40_3, L2, EM
Conflict for <variable>, <dimension>: units.	40_4, L2, EM
Conflict for <variable>, <dimension>: size.	40_5, L2, EM
Conflict for <variable>: changed name of a dimension.	40_6, L2, EM

Auxiliaries (non-target variables) **and global attributes:**

Auxiliary is missing in sub-temporal file.	50_1, L2, PT, EM
Additional auxiliary <name> in sub-temporal file.	50_2, L2, PT, EM
Auxiliary <name>: new attribute <name> for sub-temporal file.	51_1, L2, PT, EM
Auxiliary <name>: attribute <name> is missing in sub-temporal file	51_2, PT, EM
Auxiliary <name>: attribute <name>: changed value in sub-temporal file.	51_3, L2, PT, EM
Global attribute <driving_experiment>: wrong number of items.	52_1, L1
Global attribute <driving_experiment, item=num>: in conflict with attribute <name>.	52_2, L1
Auxiliary <plev | plev_bnds> for cloud amounts: unmatched required value(s)=<val>.	53_1, L2, EM
Auxiliary <plev_bnds> for cloud amounts: missing.	53_2, L2, EM
Auxiliary <name>: number of dimensions has changed.	54_1, L2, PT, EM
Auxiliary <name>: name of dimension has changed.	54_2, L2, PT, EM

Auxiliary <name>: checksum of levels or grid values has changed.   54_3, L2, PT, EM

Data:

Data set entirely of const value.	60_1, L1
Data set entirely of _FillValue.	60_2, L1
All time values with identical data.	60_3, L1
Time value before the first time-table range.	63_1, L2, EM
Time value after the last time-table range	63_2, L2, EM
Time value does not match time-table value.	63_3, L2, EM
Too many time values compared to the time-table.	63_4, L2, EM
Suspicion of fractional data range for units [%]	64_1, L1
Suspicion of percentage data range for units [1]	64_2, L1

Tables

Could not create lock-file to protect project table	67_1, L1
Could not open the CORDEX table with required attributes.	70_1, L3, ST, EM
Could not create project table.	70_2, L3, PT, EM
No path to tables; faulty config-file entry.	70_3, L3, ST, EM
Corrupt CORDEX standard table: missing column(s).	70_4, L4, ST
Frequency not found in the CORDEX standard table captions.	71, L3, ST
Project table is locked for more than 1/2 hour.	72_2, L3, EM
CORDEX domain Table 1: domain name does not match.	75_1, L3
CORDEX domain Table 1: value of Nlon|Nlat does not match.	75_2, L3
CORDEX domain Table 1: missing auxiliary <rotated_pole>.	75_3, L3
CORDEX domain Table 1: value of N.Pole lon|lat does not match.	75_4, L3

CORDEX domain Table 1: missing data for auxiliary longitude/latitude.	75_5, L3
CORDEX domain Table 1: resolution does not match.	75_6, L3
CORDEX domain Table 1: West|East|South|North boundary does not match.	75_7, L3

Miscellaneous

Indication of a renewal of data during QC.	80, L3, EM
File not in NetCDF-4 format.	81_1, L1
File not compressed.	81_2, L1
Status: apparently in progress.	82, D