

EARLINET Data On-Fly Quality Check – v1.0

Summary

- 1- Introduction..... 2
- 2- On-fly Technical Quality Check v1.0 3

V1.1

09 June 2017

1- Introduction

Data quality check procedures for EARLINET/ACTRIS database improved over the time. This document reports the automatic quality check procedures implemented on fly during the data submission phase after 23 February 2017.

The implementation of this on-fly technical quality checks will significantly reduce the number of files on the EARLINET database not compliant with EARLINET standard. A feedback to the Data Originator guarantees the possibility for data sanitization. During the uploading phase a log of the on-fly QC procedure is produced for the Data Originator detailing all the failed QC procedure for each submitted file. The on-fly QC procedure as well as the feedback mechanism work both for the single file upload and for the multiple files upload systems.

Details of the specific on-fly QC are reported in the following sections.

2- On-fly Technical Quality Check v1.0

TQC_00

This procedure checks that each file contains the mandatory product:

QC_00a	applies on b files	==> Backscatter, ErrorBackscatter IS NOT NaN or NULL (has at least 1 valid value)
QC_00b	applies on e files	==> Extinction, ErrorExtinction IS NOT NaN or NULL (has at least 1 valid value)

TQC_01

This procedure checks that the name of the file is correct respect to expected one (correct ID, Datetime and wavelength).

QC_01a check on [STATIONID [2 chars]]

QC_01b check on [DATETIME [yy][MM][DD][HH][mm]]

QC_01c check if Filename [DATETIME] is consistent with the measurement [starttime] in the database
In particular this checks if the StartTime and the name of the file are compliant. StartTime is considered compliant if it is equal to the time reported in the Filename within 2 minutes.

TQC_02

This procedure performs the following checks :

- 1 - check if file length is 17 or 18 characters
- 2 - check the 1-st and 2-nd file characters are [char]
- 3 - check the 13-th file characters is [.]
- 4 - check the 14-th file characters is [e/b]
- 5 - check the chars from 2 to 11 are decimals [datetime]
- 6 - check the chars from 14 to 16/17 are decimals [wavelength]

TQC_03

Here the procedure checks if file extension i is consistent (within 1nm) with the Emission Wavelength in the internal data table of the database, needed for searching procedure on the EARLINET database interface.

TQC_04

This QC procedure checks for Undefined Variables: it is expected that defined variables should contain at least one valid number. Further checks on MixingLayerHeight and DustLayerHeight are performed respect to their definition reported in the EARLINET file content description (<https://www.earlinet.org/index.php?id=125>)

- QC_04a All array variables cannot be ALL undefined or negatives
- QC_04b If it is a b-file must contain the variables Backscatter, ErrorBackscatter
- QC_04c If it is an e-file must contain the variables Extinction, ErrorExtinction
- QC_04d If the MixingLayerHeight is present must also be present the DustLayerHeight
- QC_04e $\text{MixingLayerHeight} \leq \text{DustLayerHeight}$
- QC_04f $\text{Altitude of the station} < \text{MixingLayerHeight}, \text{DustLayerHeight}$

This QC procedure applies also to the time series.

TQC_05

This procedure checks that the coordinates reported in the file are compliant with EARLINET station coordinates.

It checks if (latitude, longitude, altitude) of location are correct within the errors:

- (latitude, longitude) ± 0.05 degrees \Rightarrow 5 km
- altitude ± 20 m