



New data quality control procedure MetQC for operational use

Petr Stepanek (1,2), Pavel Zahradnicek (1,2), Antonello Squintu (3), and Gerard van der Schrier (3)

(1) Global Change Research Institute, Czech Academy of Sciences, Bělidla 986/4a, 603 00 Brno, Czech Republic, (2) Czech Hydrometeorological Institute, Brno Regional Office, Kroftova 43, 616 67 Brno, Czech Republic, (3) KNMI, De Bilt, Netherlands

Preparing data quality control software for operational task is time consuming and hard programming task. The quality control itself must be able to capture the most typical errors that occur in climatological data but also to detect specific cases. The current software to be presented here is based on years of experience in testing various datasets.

Within the context of Copernicus project C3S.311a.Lot.4, where CzechGlobe (GCRI) and KNMI take part, a new procedure and software (MetQC) based on R was developed and applied in operational mode.

The QC procedure is based on best ideas of originally stand-alone software (ProClimDB, MASH, ACMANT, NOAA, C3QC). After the methods comparison and evaluation, based on benchmark dataset utilizing both real and surrogate data, a new R script was programmed. These scripts will be available for the scientific community for further open development.

The software has been already tested in operational mode at Czech Hydrometeorological Institute (daily and sub-daily data of various meteorological elements) and also at KNMI on ECA&D (daily dataset). On the output, it gives list of problematic values together with a lot of auxiliary information including graphical outputs etc. Duplicated stations and repeated values are revealed and reported as well.

The MetQC includes these individual parts: calculation of correlations, finding neighbours for subsequent data quality control and data quality control itself. Data quality control is carried out by combination of several approaches, based on both time and spatial analysis. With proper settings for a particular meteorological element, it is possible to automate the process. Even if the process is automated, the software still gives users full control about the process – e.g. by possibility to check auxiliary outputs recorded with full detail during the calculations.