QC/QA flagging ad hoc group progress report (HP3)

QTF team,
assembled by Tzvetan Simeonov
GRUAN Lead Centre, DWD

13th GRUAN Implementation and Coordination Meeting (ICM-13)

Virtual

Session 1, 15 November 2021

Status from last presentation in ICM-11

- RS92-GDP.2 → some rules are defined by LC → incl. assessment
- RS-11G-GDP.1 → other rules are defined by PC → NO assessment
- Until now → NO GRUAN-wide scheme and NO general rules for controlling & assessment are defined yet

- The Lead Centre should not dictate such rules!
  - Some ideas are available, but more are necessary
  - Discussion is really needed with interested persons from community
  - A formal definition of rules is required (published as TN or part of TD)

- Wanted! → Group of persons to
  - think about general scheme and details → discuss and decide
  - write down and publish as official GRUAN document

List of discussed issues

- Issue 1: gap size threshold / data interpolation
- Issue 2: parameter-dependent gap treatment
- Issue 3: uncertainties treatment of interpolated values
- Issue 4: flagging missing values
- Issue 5: flagging for cloud type of profile

Issues 2&4: flagging missing values

Complete agreement upon flags:

- **Each point in the profile** should be flagged individually for missing data or interpolated data if any

Complete agreement upon variables:

- **Each variable** should be flagged separately

Issue 1: gap treatment in data

Proposal of the group majority:

- Interpolate missing values in gaps, smaller than 10s
- Flag interpolated gaps for every parameter individually

Opinions in favor of interpolation:

- 95% of soundings have gaps of less than 10s
- All gaps are flagged, so users can ignore interpolated data
- Treatment of interpolation and flags should be variable-dependent

Opinions in against interpolation:

- GDP's are observation products, not interpolation products

Issue 3: uncertainties of interpolated data

Agreement on enhanced uncertainty among the proponents of interpolation of data

Disagreement with uncertainties among opponents of interpolations – how can we calculate uncertainties for data we have not collected?

Approach 1:

- Calculate uncertainties from mean uncertainties values by enhancing the uncertainty with a Gaussian process

Approach 2:

- Calculate uncertainties enhancement from a look-up table, such as in *Fasso et. al. 2020*

Interpolation uncertainty of atmospheric temperature radiosoundings

Alessandro Fassó[1], Michael Sommer[2], and Christoph von Rohden[2]

[1]University of Bergamo, Italy
[2]GRUAN Lead Centre, Deutscher Wetterdienst, Lindenberg, Germany

Correspondence: Alessandro Fassó (alessandro.fasso@unibg.it)

Abstract. This paper is motivated by the fact that, although temperature readings made by Vaisala RS41 radiosondes at GRUAN sites (www.gruan.org) are given at 1 s resolution, for various reasons, missing data are spread along the atmospheric profile. Such a problem is quite common in radiosonde data and other profile data. Hence, (linear) interpolation is often used to fill the gaps in published data products. In this perspective, the present paper considers interpolation uncertainty. To do this, a statistical approach is introduced giving some understanding of the consequences of substituting missing data by interpolated ones.

Issue 5: flagging clouds

No reliable cloud detection from the radiosonde sensors

Cloud observations available at launch from SYNOP data

Advantages on using global flag:

- Availability of general flag for classification
- Using already available data from the observation procedure
- Low clouds very relevant

Disadvantages on using global flag:

- Cloud conditions are changing constantly at launch site
- Clouds change throughout the profile
- Medium and high clouds observations unreliable

RS41GDP implementation

Gaps up to 10s interpolated

Enhanced uncertainties for interpolated points

Flags for every point:

- *dorn* – Day or night
- *band* – specify radiosonde state period – gc/ascent/descent
- *parameter_raw_flag* – individual flags for every measured parameter (nan, outlier, interpolated, out of range, zero)

Global flags:

- GC availability
- SHC checks
- SYNOP observations, including cloud types
- Product type – GDP versions
- Daytime – daytime/nighttime/twilight

QTF next steps

- Implementation of the discussed QC/QA universally in all current and future GDP's
- Universal GPD parameter name conventions

Thanks to QTF team

- Live opinion polling among the community
- Gratitude for contributions from Alessandro, Christoph, Dale, Frederic, Gianni, June, Masatomo, Michael, Shunsuke and Tony