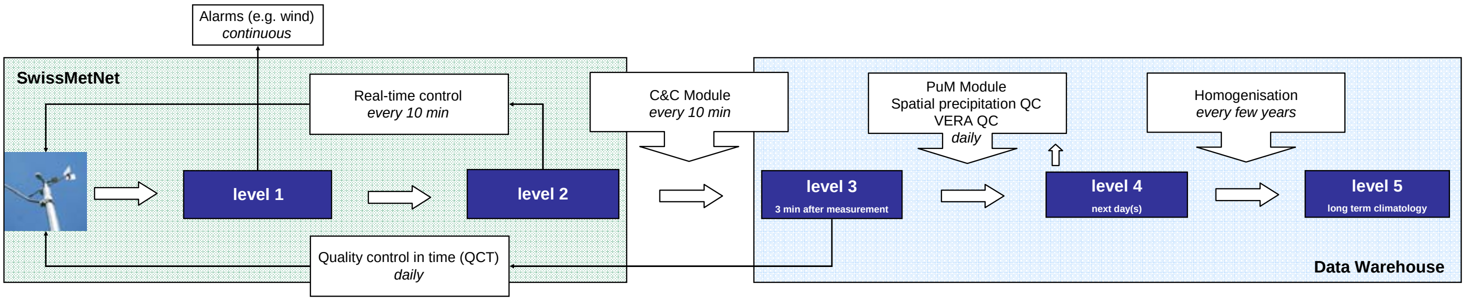
The data quality control chain for automatic surface observation data at MeteoSwiss

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Home Affairs FDHA
Federal Office of Meteorology and Climatology
MeteoSwiss

D. van Geijtenbeek, C. Naguel, S. C. Scherrer, B. Landl
Federal Office of Meteorology and Climatology (MeteoSwiss), Switzerland



Technical survey – detecting of sensor troubles

The technical test chain is designed for detecting instrument failures and sensor troubles based on house keeping quantities combined with climatological parameters. The technical part of quality control allows the monitoring of ground-based stations and has no effect to the data series.

Onsite instantaneous alarms
Each meteorological station has a local Automatic Data Acquisition System (ADAS) which among others delivers instantaneous alarms (e.g. windalarms or alarms for surface freezing) directly to the clients.

Real-time control (plausibility tests online)
In the Central Data Acquisition System (CDAS) real-time controls are performed, using an integrated quality control with threshold values and dead band criterias to detect instrument failures. These online plausibility tests deliver instantaneous alarms and warnings to a network surveying center.

Fig. 1: SwissMetNet – the new automatic meteorological ground-based network of Switzerland

Quality control in time (QCT) on raw data
See presentation 4-O3; Barbara Landl

The Quality Control in Time (QCT) is operationally run on a daily basis (using measured raw data from the previous three months).

This quality control allows detecting drifting time series due to instrumental problems which can not be seen by real-time controls. Thanks to this control the time to detect instrumental problems can be reduced and therefore the maintenance and on-site interventions can be optimized. Furthermore this results in an improvement of the measurement accuracy and of the data availability.

References

Steinacker, R., Ch. Häberli, W. Pöttschacher, 2000: A transparent method for the analysis and quality evaluation of irregularly distributed and noisy observational data. *Mon. Wea. Rev.* **128**, 2303-2316.

Frei, C. and Schär, C. 1998: A precipitation climatology of the Alps from high-resolution rain-gauge observations. *Int. J. Climatol.* **18**, 873-900.

Contact

deborah.vangeijtenbeek@meteoswiss.ch
claudine.naguel@meteoswiss.ch
simon.scherrer@meteoswiss.ch
barbara.landl@meteoswiss.ch

Climatological survey – flaging of suspicious data / removal or correction of faulty and missing data

The tests for climatological survey are classified into four main categories (according the recommendations of WMO). Their goal is the cleaning and enhancing of climatological time series.

- **Limit tests.** Most variables are compared to physical ('hard'; e.g. 0 and 360 deg for wind direction) and climatological ('soft') limits. For the climatological limits, the 99.9% percentiles were determined for each month and station from 15-20 year time series.
- **Variability tests.** There are two different types of variability tests: one which tests the maximally allowed variability during a specified time interval and one which tests the minimal required variability during a certain period ('dead band' range).
- **Inter-parameter consistency tests.** Values measured at the same time and at the same place may not be inconsistent to each other (e.g. 8 octas of total cloud cover with bright sunshine).
- **Spatial consistency tests.** Values of the same parameter measured at the same time at nearby stations may not differ too much.

Inter-parameter integrity

At MeteoSwiss the inter-parameter integrity is checked at high temporal resolution (10 min). The chain is divided into two deferred modules, C&C and PuM.

In order to assure that both modules use the same set of tests and constants, the control data is implemented as the core of the QC system. At present about 220 inter-parameter integrity tests are in operation. The modules use a unified flagging procedure consisting of the 'plausibility information' (a bitmap indicating the test violations) and a 'treatment information' (indicating the correction if one was applied). The web based application for submission of visual observations is also linked to this system. The system returns quality check information directly to the observer while transmission.

C&C Module: Calc and Check
The Calc and Check module act without user interaction. It flags suspicious data and deletes obvious errors automatically before entering the data warehouse every ten minutes. This module also calculates the derived quantities at 10 min resolution. Three minutes after the measurement has been completed, the checked and calculated values are available in the central database. The C&C Module ensures a minimal quality level for real time data purposes.

PuM Module: Plausibility and mutation
This process starts automatically once a day and is split into a part of automatically treatment and a second part which allows an interactive checking and manual mutation of data for stations with a high climatological importance.

In a first step, the data are tested in a similar way as in the C&C Module but including tests for temporal variability. In a next step PuM automatically generate replacement values, if there are short gaps in the data (6h at maximum) with mean rational spline method (Fig. 4). Interpolation cases, which can not be handled automatically, as well as suspicious values which violated any integrity test, are handed over to the interactive processing.

Fig. 2: GUI for interactive checking and mutation (e.g. soft limit violation)

The tool for interactive processing has a graphical interface which presents the data on a spreadsheet and as time series plots (Fig. 2). At any stage the operator can test the corrections and confirmations. If there are no more tests violated, the case is considered as being solved.

Spatial integrity

The spatial integrity tests use spatio(-temporal) information to find suspicious data and to provide objective interpolation values. At MeteoSwiss two different methods are in use for spatial quality checks.

Spatial precipitation QC
See presentation 4-O4; Simon C. Scherrer

Based on the SYMAP gridding algorithm for precipitation (Frei and Schär, 1998) an objective method to detect suspicious daily sums is implemented. One of the main challenges is the spatial interpolation in complex terrain of the alpine region. The spatial precipitation QC method helps to objectify treatment of precipitation data and to reduce the workload for the data editors (incl. splitting of several day measurements). They can concentrate on editing outliers instead of finding them.

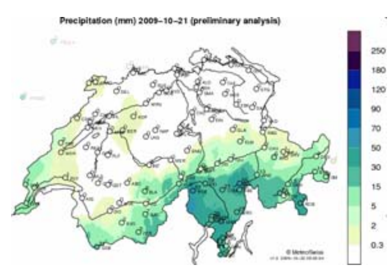


Fig. 3: Spatial precipitation QC is based on the SYMAP gridding.

VERA QC
The VERA QC Module was developed at the Department of Meteorology and Geophysics at the University of Vienna (Steinacker et al., 2000). The VERA spatial interpolation algorithm is used to generate a virtual comparative measurement (hourly values for temperature, humidity, pressure and wind components). This virtual parameter can be easily integrated in the existing PuM infrastructure and can be used for plausibility tests as well as for improvement of the automatical interpolation (Fig. 4). In near future VERA QC will be integrated into the operational QC chain.

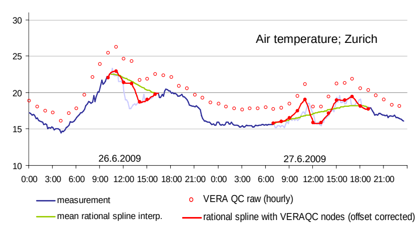


Fig. 4: Comparison between different interpolation modes (gap 6h and 12h)