Summary of three quality control algorithms for rainfall data from personal weather stations

Jochen Seidel, Lotte de Vos, Tess O‘Hara

PWS – What‘s out there

- Netatmo PWS are the dominant low-cost PWS in many regions (e.g. Germany, France), but there are also other types of PWS and PWS networks such as e.g. WOW, Weather Underground, etc.
- Wide range of devices ranging from low cost PWS (Netatmo) to „semi-professional“ stations (e.g. Davis)
- Large number of PWS >> Gauges from national weather services
- PWS may provide reliable rainfall data **if** these stations are set up and maintained **correctly…**

PWS – Setup and errors

- Installation conditions of PWS are unknown (mostly not according WMO standards)
- Maintainers of PWS can be anybody from absolute “beginners” with no meteorological background to hobby meteorologists
- Especially Netatmo PWS suffer from typical errors (faulty zeroes, high influxes, wrong locations, …..)

Which QC method(s) is/are suitable for which PWS data set(s)?

QC Modules

PWSQC	PWSpy-QC	GSDR-QC-local
1. Faulty zeroes filter	1. Indicator based filter	1. Flagging of suspicious observations using defined rule base
2. High influx filter	2. Bias correction	2. Filtering of suspicious observations not meeting QC criteria
3. Station outliers filter	3. Event based filter	
4. Bias correction		

Level of QC allocation

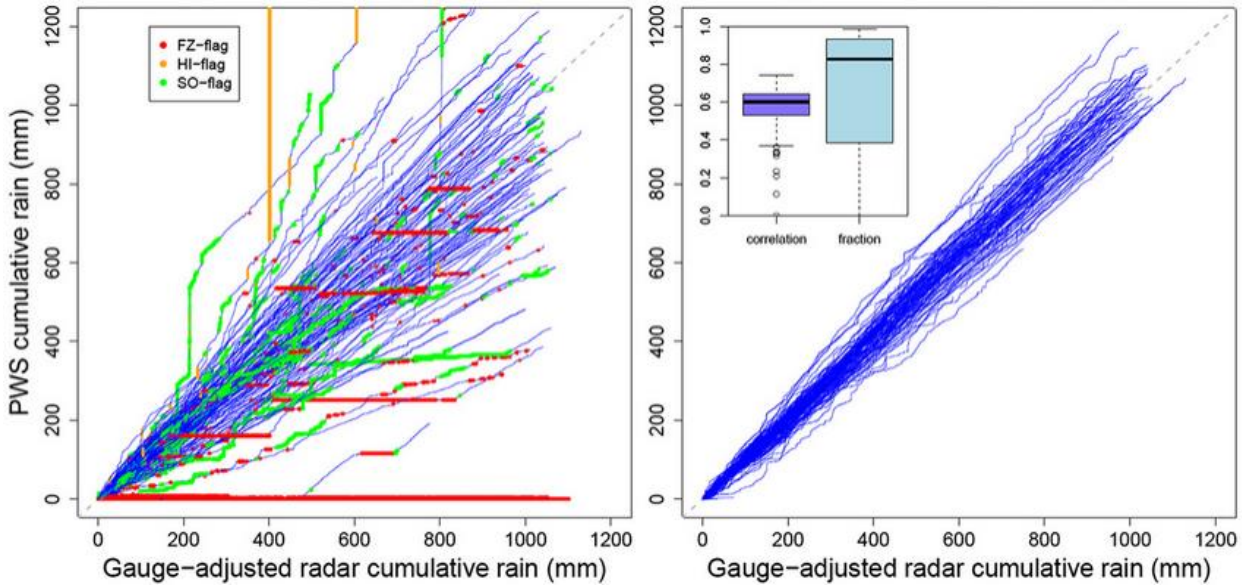
PWSQC	PWSpy-QC	GSDR-QC-local
Per individual measurement	Per full PWS time series	Per individual measurement
Dynamic nature is suitable for longer time series	Event based	Dynamic nature is suitable for longer time series

QC Output

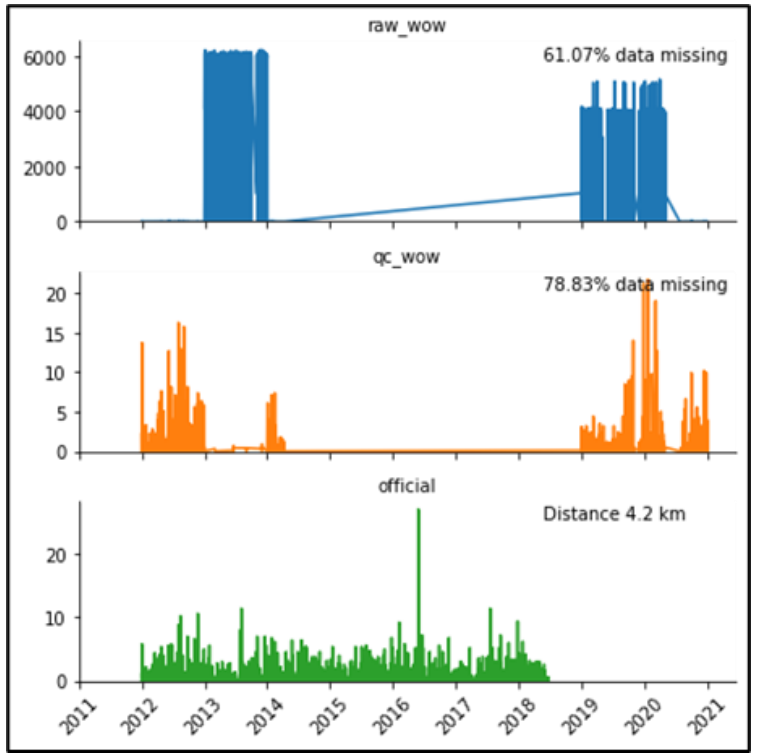
PWSQC	PWSpy-QC	GSDR-QC-local
Original PWS dataset	Set of trustworthy PWS	Flag file for each PWS gauge showing individual test results
3 flag files conveying flag attribution to individual observations for QC modules	Individual bias correction for each time series	Output file with reliable observations
1 file with Bias correction factors generated for each observation	Implausible time intervals removed for each accepted time series	
Bias adjusted PWS dataset with only reliable observations		

Examples for QC Output

PWSQC



GSDR-QC-local



Recommendations for PWS QC Methods

Which QC method(s) is/are suitable for which PWS data set(s)?

PWSQC	PWSpy-QC	GSDR-QC-local
Developed for (Netatmo) PWS	Developed for (Netatmo) PWS	Developed for reliable PWS
„Bulk“ bias correction factor	Individual PWS Bias correction	No Bias correction, less sensitive to faulty zeroes
No reference data set required	Trustworthy reference data set required	Trustworthy reference data set required
Flags suspicious data	Rejects complete time PWS series	Flags/removes suspicious data
Dense PWS network	Dense PWS network	Sparse and heterogeneuos PWS
5 minute data	Hourly data	Sub-daily time intervals

Outlook / Next Steps

- Improving the algorithms (depending on purpose)
  - retaining more or more reliable PWS
  - infilling of data gaps in PWS
  - identifying high influxes correctly
  - speeding up calculations
  - real time application
- Combining PWS QC methods/modules from different QC methods
- Merging with other OS data (e.g. CML)
- Using PWS for other applications such as radar gauge adjustment
- Communicating the benefit of imperfect data
- Evaluating the need of potential users to find further collaborations