A toolbox for real time data acquisition and quality control of personal weather station rainfall data

Jochen Seidel (1), Georges Schutz (2) and Nathalie Rombeek (3)

[1] Institute for Modelling Hydraulic and Environmental Systems, University of Stuttgart, Germany, [2] RTC4Water s.à.r.l., Roeser, Luxembourg, [3] Department of Water Management, Delft University of Technology, the Netherlands

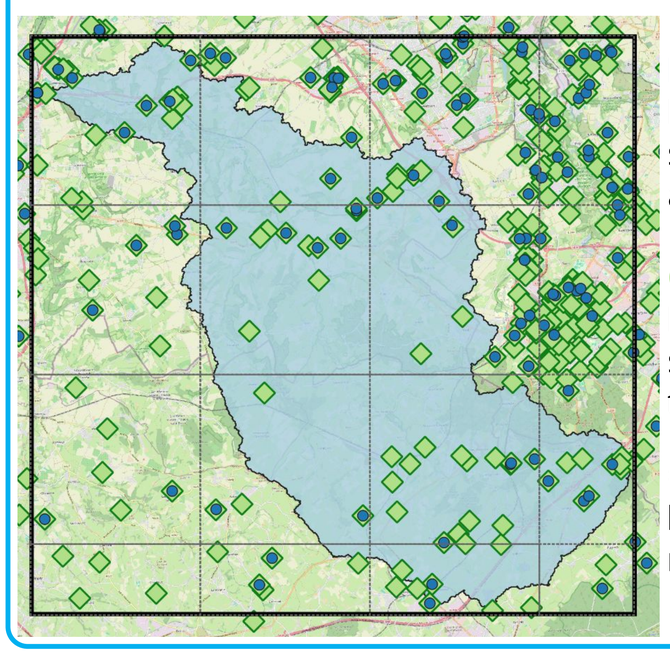
Introduction

- Personal weather stations (PWSs) have often a **higher network density** compared to official weather stations from national meteorological agencies and are **available in near real-time**, offers a large potential to improve precipitation estimates and can potentially be used for nowcasting, flood forecasting or early warning system.
- **Latency** and **quality** of the data are important aspects to consider.
- We explored the **real-time potential** of rainfall data from PWSs from the private company Netatmo, which can be accessed via Netatmo's application programming interface (API).
- We developed a **user-friendly toolbox** for downloading PWS data using Netatmo's API.
- First results of these analyses are presented.

Scanning strategy

Different strategies for station scanning:

a) Data scraping from web page with different zoom level (does not require an account, however, this procedure time consuming).
b) Using the API [1] for small tiles within a given area.



Blue circles: PWSs obtained by scanning the large area.

Green diamonds: PWSs obtained by scanning smaller tiles.

⇒ Area size might lead to different results!

Aspects of (real-time) data acquisition

- Temporal resolution ~ 5 min.
- Resampling issues.

PWS 1 time (raw)	PWS 2 time (raw)	Resampled time from Netatmo	Comment
05:49:40	05:49:23	05:50:00	
05:54:48	05:54:20	05:55:00	More or less representing the same time interval
05:59:53	-	06:00:00	
06:04:40	06:00:03	06:05:00	Not representing the same interval

- Data contains number of tips within a timestamp, the time of tips is not available
- update cycle of ~ 10 min.
  - Irregular data latency of max 10 min.
  - 0 - 2 Datasets in that time period.

Request time	PWS 1	PWS 2	PWS 3	PWS 4	PWS 5	PWS 6	PWS 7	PWS 8	PWS 9	PWS 10
10:02:30	Y		X	X	X			X		Y
10:05:00		Y				X	Y		Y	
10:07:30	X		Y	Y	Y			Y		X
10:10:00		X					X		X	
10:12:30			X	X	X			X		

Quality control aspects of real-time Netatmo PWS data

- Quality control from De Vos et al. [2].
  - Faulty zero and high influx filter can be potentially used in real time, small changes required, work in progress.
  - Station outlier filter should be applied once per time interval to check is if a specific PWS is reliable or has become an outlier, e.g. by the owner changing PWS location.
- Indicator Correlation (IC) from Bárdossy et al. [3] requires a certain data length. However, an IC-filter with a monthly rolling window could be possible to assess if PWS are presumably reliable.

⇒ However, individual PWSs can potentially have issues at a give time step.

⇒ Flagging such PWSs might be helpful.

Toolbox

Currently developed on Github, includes routines for:

- Scanning PWSs in a specified area (shapefile or bounding box).
  - Checks at which time (YYYY-MM) a PWS started recording.
- Downloading station data for all variables (rainfall, temperature, humidity, pressure and wind).
- Combining downloaded data to longer time series.

Limitations

- Limited number of requests possible using the API.
- Available PWSs changes over time, list needs to be updated.
- Scanning large areas only provides a random subset of PWSs within that area.

Some examples of real-time application

- Waterboard Rijn en IJssel in the Netherlands [4].
- MET Norway uses Netatmo PWSs for improving weather forecasts.
- EUMETNET contract with Netatmo for 6 months in summer 2025, distributes real-time data from ~300.000 PWSs to its' partners.

Outlook

- Publish toolbox on OpenSense Github.
- Optimizing real-time data acquisition.
- Include tools for resampling data.
- Check for other PWS sources, e.g. WOW, MeteoNetwork, WeatherUnderground.

Literature

[1] NETATMO: Welcome aboard!, available at: https://dev.netatmo.com/apidocumentation, last access: 10 June 2025.

[2] De Vos, L.W., Leijnse, H., Overeem, A., and Uijlenhoet, R.: Quality control for crowdsourced personal weather stations to enable operational rainfall monitoring, Geophysical Research Letters, 46, 8820–8829, 2019.

[3] Bárdossy, A., Seidel, J., and El Hachem, A.: The use of personal weather station observations to improve precipitation estimation and interpolation, Hydrol. Earth Syst. Sci., 25, 583–601, https://doi.org/10.5194/hess-25-583-2021, 2021.

[4] https://neerslagdata.wrij.nl/inzicht, last access: 10 June 2025.