COST Action CA20136
Opportunistic precipitation sensing network - OPENSENSE

D3.1 Opportunistic sensing precipitation product validated using traditional observations and published as a dataset

Version 1.0

Working Group	WG3
Deliverable	D3.1
Date	20/05/2025
Version	1.0
Author	Maximilian Graf, Remco van de Beek
Reviewer	Martin Fencl

Description	This document describes the validation of OS datasets and where they can be found
Key words	OS data, validation, dataset

About OPENSENSE (COST Action CA20136). *OPENSENSE brings together scientists investigating different opportunistic sensors (e.g. microwave links, citizen science), experts from weather services, and end-users of rainfall products to build a worldwide reference opportunistic sensing community. The overarching goals of the COST are to overcome key barriers preventing data exchange and acceptance as hydrometeorological observations, define standards to allow for large-scale benchmarking of opportunistic sensing precipitation products and develop new methods for precipitation retrieval, coordinate integration of the opportunistic observations into traditional monitoring networks, and identify potential new sources of precipitation observations. Further details can be found here:*

Funded by the European Union

www.cost.eu
www.opensenseaction.eu

COST (European Cooperation in Science and Technology) is a funding agency for research and innovation networks. Our Actions help connect research initiatives across Europe and enable scientists to grow their ideas by sharing them with their peers. This boosts their research, career and innovation.

Table of contents

Funded by the European Union

Glossary

WG	Working Group
OS	Opportunistic Sensors

1. Introduction

The deliverable D3.1, Opportunistic sensing precipitation product validated using traditional observations and published as a dataset, aims to validate the efforts of WG1's data collection and WG2's data processing activities. For both CMLs and PWS, OpenSense members are conducting benchmarks of processing and quality control methods with open source software compiled by WG2 using an open dataset gathered by WG1. Additionally, merging studies are conducted using open radar and CML datasets, using CML and PWS data for radar adjustment, as well as a pan-European radar merging with PWS.

2. Validated products

2.1. Benchmarking CML processing methods

Graf et al. (to be submitted, 2025) benchmark various processing methods for rainfall estimation using data from two open commercial microwave links (CMLs) datasets, which are part of the existing telecommunication infrastructure. Namely, these datasets are the openMRG (Andersson et al. 2022) and the OpenRainER dataset (Covi et al. 2025). CMLs offer valuable opportunities for opportunistic environmental monitoring, especially in regions lacking dense hydrometeorological networks. The authors compare multiple CML processing software packages, focusing on performance, accuracy, and applicability. The benchmark is conducted using two openly available datasets, allowing an objective assessment of each method's ability to retrieve path-averaged rainfall intensities. Key metrics such as sensitivity to wet antenna effects, temporal resolution handling, and retrieval accuracy are analysed. The study emphasises the trade-offs between real-time processing capabilities and retrieval robustness under different environmental conditions. Results indicate that while some tools excel in high-resolution and real-time applications, others perform better under complex environmental noise scenarios. A special emphasis is placed on the reproducibility and transparency of the processing chain, recommending open-source solutions where feasible. The results are validated using rain gauge and radar data that are available together with the CML data. The paper concludes that there is no universally optimal method, and the choice should depend on specific operational goals, data availability, and environmental context. The findings are intended to guide both researchers and practitioners in selecting appropriate CML processing strategies for rainfall estimation tasks. An example of the validation of the two CML datasets is shown in Fig. 1. The datasets are already publicly available, and the reproducible code to derive the results is available under https://github.com/OpenSenseAction/radar_adjustment_intercomparison.
This study will be submitted during GP 4 of OpenSense.

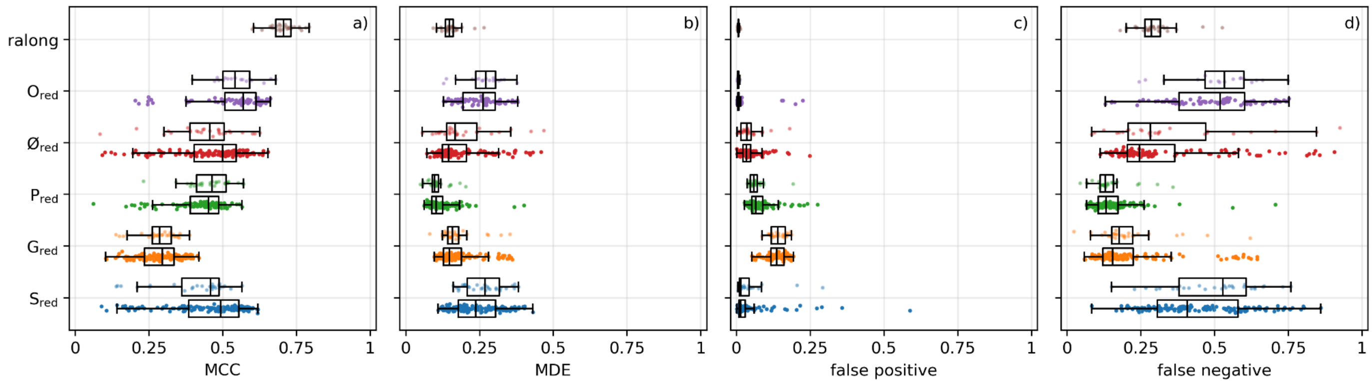

Fig 1. Binary evaluation of five rain event detection methods for the OpenRainER dataset. Each colour indicates one method. The upper, lighter points and boxplot show the metric for rain gauges as a reference, and the lower, darker points and boxplot show the metric for ralong (radar along the CML path) as a reference. Additionally, the rain gauges used as reference for the CMLs (within two km) are indicated by brown points.

2.2. Comparing PWS QC methods

Similar to the approach in the comparison of CML processing methods in 2.1, we are conducting a comprehensive evaluation of quality control (QC) techniques for personal weather stations (PWS). This comparison is based on the open-source package pypwsqc (https://github.com/OpenSenseAction/pypwsqc) and hosted in a dedicated, publicly accessible repository: https://github.com/OpenSenseAction/pws_qc_intercomparison. The goal is to assess the performance and reliability of various QC methods implemented in *pypwsqc*, which incorporates algorithms derived from several peer-reviewed publications.

For the dataset, we utilise openly available observations from PWS networks, including a high-resolution dataset from Amsterdam, originally published by de Vos et al. (citation pending), and potentially also from the EUMETNET PWS Sandbox dataset. Within the *pypwsqc* framework, the Amsterdam dataset is processed and validated against a reference radar rainfall product to evaluate the effectiveness of different QC approaches in filtering noise and correcting biases.

This data comparison formed a part of the second OpenSense Training School, held in April 2025 in Budapest, where participants engaged in practical exercises using this dataset and QC workflows. Materials from that event, including an introductory notebook that validates the available QC methods for the Amsterdam dataset and its reference data, are available online:

https://github.com/OpenSenseAction/TrainingSchoolMergingApplication/blob/main/intro_session/3_pws_qc_new.ipynb.

2.3. Merging radar with CML data

A small team of WG2 of QG and WG3 members are conducting an intercomparison study on CML-radar merging techniques based on two open CML datasets. The group consists of Erlend Oydvin, Elia Covi, Maximilian Graf and Christian Chwala.
They use CML and radar from the openMRG dataset covering Gothenburg (Andersson et al. 2022), Sweden and the OpenRainER database (Covi et al. 2025), covering the Emilia-Romagna region, Italy. They compare additive and multiplicative adjustment methods as well as kriging with external drift. The adjustment methods, that can take the line characteristic of the CMLs into account, are compiled in the open Python software package (https://github.com/OpenSenseAction/mergeplg).
The merged rainfall maps are validated against rain gauges that are available in both open datasets. All code and results will be available online at https://github.com/OpenSenseAction/radar_adjustment_intercomparison, and an example is given in Fig. 2 showing the improvement of different adjustment methods (x-axis) compared to reference rain gauges. We expect the study to be submitted within GP4 of OpenSense.

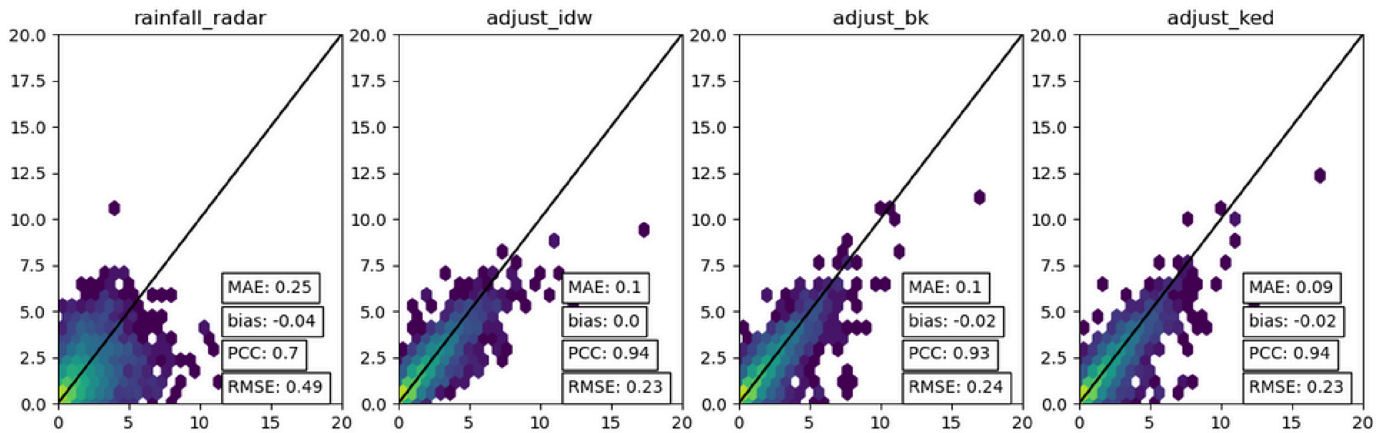

Fig. 2. Radar rainfall and three CML adjusted rainfall products (x-axis) compared to rain gauges (y-axis) for the OpenMRG dataset. The methods are additive adjustment using inverse distance weighting, additive adjustment using block kriging and kriging with external drift.

2.4. Danish CML and PWS data merged with radar data

The study by Nielsen et al. (2024) shows for a part of Denmark that the combination of weather radar data with opportunistic precipitation sensors can significantly improve precipitation estimates. Data from CMLs and private weather stations (PWS) were used. The researchers developed a data fusion approach called “Moving Median Bias Adjustment”. This approach made it possible to create a high-quality precipitation product that works independently of conventional official rain gauges. The results showed that the merged data achieved a Nash-Sutcliffe efficiency (NSE) of up to 0.88 for daily accumulated precipitation, while the individual products had NSE values between -7.44 and 0.65. The datasets for this study were not published.

2.5. PWS and a pan-European radar dataset

OpenSense partners, the department R&D Observations and Data technology from the Royal Netherlands Meteorological Institute, have published a paper on 14 February 2024 in the Hydrology and Earth System Sciences journal. This paper is available in open access: Overeem, A., Leijnse, H., van der Schrier, G., van den Besselaar, E., Garcia-Marti, I., and de Vos, L. W.: Merging with crowdsourced rain gauge data improves pan-European radar precipitation estimates, Hydrol. Earth Syst. Sci., 28, 649–668, https://doi.org/10.5194/hess-28-649-2024, 2024.

An important potential use case for opportunistic sensing precipitation sensors is to improve radar precipitation products. Ground-based radar precipitation products provide wide coverage with high spatial resolution, but typically need to be adjusted with rain gauge accumulations to achieve sufficient accuracy. However, rainfall events can be missed by the limited spatial density of (near) real-time rain gauges from official networks. Through their much higher spatial density, crowdsourced rain gauges are a potentially interesting complementary source for improving radar precipitation products. Here, a 1-year crowdsourced personal weather station (PWS) ~5 min rain gauge dataset has been acquired from the private company Netatmo covering Europe over the period 1 September 2019–31 August 2020. Quality control was applied to the PWS rain gauge data using neighbouring PWSs and unadjusted radar accumulations. Finally, 1-h PWS gauge accumulations are merged with 1-h European OPERA radar accumulations. The obtained merged radar-PWS precipitation dataset is publicly available.

Evaluation of the merged radar-PWS 1-h and daily accumulations against independent official rain gauges available through the European Climate Assessment & Dataset (ECA&D) confirms the potential of crowdsourced PWS rain gauge data for improving radar precipitation products in (near) real time. The severe mean underestimation for daily precipitation of ~ 28 % from the unadjusted radar dataset (OPERA) decreases to ~ 3 % for the merged radar–PWS dataset. A spatial verification at official rain gauge locations shows that the relative bias is still quite variable, and underestimation seems more severe for colder climates. Scatter density plots of radar-PWS versus official gauge accumulations for different temperature classes reveal that the performance is indeed less good for lower temperatures. This points to limitations in observing solid precipitation by PWS gauges.

Dataset: Overeem, A. (Aart); Leijnse, H. (Hidde); van der Schrier, Gerard; van den Besselaar, Else; Garcia-Marti, Irene et. al. (2024): Dataset of pan-European 1-h OPERA radar precipitation accumulations adjusted with rain gauge accumulations from Netatmo personal weather stations. Version 2. 4TU.ResearchData. dataset. https://doi.org/10.4121/675f3f64-04a8-48db-ae3e-4a6c004a0776.v2

References

Andersson, J. C., Olsson, J., Van de Beek, R., & Hansryd, J. (2022). OpenMRG: Open data from Microwave links, Radar, and Gauges for rainfall quantification in Gothenburg, Sweden. *Earth System Science Data*, *14*(12), 5411-5426.

Covi, E., & Roversi, G. (2024). *OpenRainER, Zenodo [data set];* https://zenodo.org/records/14731404.