D1.6 Documentation for accessing operational OS data

Version 1.0

Working Group	WG1
Deliverable	D1.6
Date	13/10/2025
Version	1.0
Authors	Roberto Nebuloni (CNR-Italy) and Martin Fencl (CTU-Czech Rep.)
Reviewer	Vojtěch Bareš (CTU-Czech Rep.)

Description	This document describes several successful use cases of real-time or near-real time access to opportunistic data (CMLs, SMLs or PWS data). Most of them were implemented within national or international research projects involving members of the OpenSense COST Action.
Key words	Datasets, OpenSense, Opportunistic data, Operational access, Precipitation, Real time, CML, SML, PWS

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*

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.

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:

Table of contents

Funded by the European Union

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.

Glossary

CML	Commercial Microwave Links
csv	Comma-separated values (filename extension associated with text files)
DM.N	Deliverable N of working group M
GPK	Grant period K
MoU	Memorandum of Understanding
NDA	Non-Disclosure Agreement
nc	Filename extension of the NetCDF (Network Common Data Form) data format
OpenSense	Opportunistic precipitation sensing network
OS	Opportunistic Sensors
PWS	Personal Weather Stations
SML	Satellite Microwave Links
SNR	Signal-to-Noise Ratio
WG	Working Group

1. Introduction

This document is the official OpenSense deliverable D1.6 *Documentation for accessing operational OS data.* It is an output of Activity 3 of WG1 (*Establishing operational access to OS*). Details about WG1 activities, milestones and deliverables due are reported in the OpenSense MoU [MoU] and listed in Table 1 as well.

Activity 3 contributes to addressing RCO6, which is one of the research coordination objectives of OpenSense. RCOs address relevant challenges in OS, in particular RCO6, sets-up the ambitious target to *Establish operational access to OS precipitation products at least in two regions to facilitate their use (and indirect verification) in precipitation and hydrological forecasts*. According to the MoU [MoU], the achievement of RCO6 is conceived as a result of the joint activities of WG1 and WG3 (Merging and Application), the latter being responsible for RC06. The following tasks contribute to RCO6:

- WG1
  - T1.3 Establishing and maintaining operational access to OS data from at least two regions/countries.
- WG3
  - T3.1. Coordinating research agenda on merging OS and standard observations.
  - T3.4. Coordinating creation of operational OS precipitation products.

The core activity of WG1 is to increase the maturity of OS data, enhancing their usage, primarily for research purposes. This work has been done through by standardization efforts, e.g. on data formats and metadata requirements (Fencl et al., 2023), by unlocking datasets and making them available on open platforms, or again, by compiling a list of benchmark datasets suitable, for instance, for testing algorithms or hydrological models. This activity also allowed us to understand the issues we are facing with OS, when we are moving from research to the operational stage. For instance, including OS data into operational data processing chains for applications in meteorology, or hydrology, requires that data streams are available to the users (e.g. meteorological offices) in real-time or with acceptable latency.

Three types of OS data have been considered so far by the OpenSense community: CMLs, SMLs and PWS. A survey of the state-of-the-art of OS usage within the OpenSense community, led to the following conclusion: at the moment, the availability of data suitable for operational purposes is limited to the PWS case, while there are no active operational processing chains exploiting either CML or SML data. This depends on the difference in the way OS data are generated, collected and transferred in the above cases as well as on data ownership.

- In the PWS case, the data are collected by meteorological stations owned by citizens. Data from many PWS can be easily transferred to a central server through the home internet connection or through mobile networks. In fact, data from many users are collected by online platforms hosted by private companies who offer meteorological services or by no-profit entities. Hence, the cost of the OS infrastructure is covered by the citizens, basically on a volunteer basis.
- In the CML case, data owners are mobile network operators and the data are primarily generated for link and network quality control. Data are stored only for the time deemed necessary to the above scope. CML raw data have been provided offline for research purposes, sometimes at a cost and/or under an NDA, as the metadata (link location and characteristics) are classified as confidential information by some operators. Apart from the technical issues in getting data from thousands of links for operators at national scale, with characteristics suitable to retrieve precipitation, in the CML case, unlocking the data in real-time or near-real time is the biggest hurdle, unless the operator can cover costs by implementing a business model.
- There are few SML cases. Again, here a business model is required to cover the costs of installing ground sensors, transferring the data to a central server and processing them. Alternatively, existing ground terminals that provide a spectrum of services (typically IoT services, such as domotics, or bank transactions in remote areas) can be exploited using the return channel (i.e. ground-to-satellite) to vehiculate the raw data required for rainfall estimation. In this case, the satellite operator (or the service provider) is the owner of the data.

In most cases, CML data were provided either for free under agreement for research use or purchased covering the costs with the budget of research projects. Similarly, SML terminals were often financed by research projects. In one case, these terminals have been deployed by a private company (HD Rain, located in France) as a part of the meteorological services provided for applications such as nowcasting or insurance against extreme events in agriculture.

Table 1 Timetable of WG1 activities, milestones and deliverables according to OpenSense MoU [MoU].

	year 1				year 2				year 3				year 4			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
WG1 - data management and standardization																
Defining common data and metadata format for data exchange				D1.1												
Sharing and curration of individual datasets				**D1.2**		**D1.3**										
Gathering large-scale OS dataset												D1.4				
Complementing of OS datasets with standard observations																
Compiling joint benchmark datasets								**D1.5**								
Establishing + maintaining operational access to OS data													**D1.6**			
Project Milestones				M1		M2		M3		M5			M6			
								M4								
WG deliverables																

WG1-related milestones

M2 Large-scale dataset in standardized format available for benchmarking of algorithms
M5 -Transboundary OS precipitation product available
M6 -Operational access to subset of OS data established

WG1 deliverables

D1.1 Repository for individual OS datasets
D1.2 White paper on data standards/formats for investigated types of OS sensors
D1.3 Documentation of past datasets shared in the repository
D1.4 Large-scale OS dataset completed by standard observations + report
D1.5 Benchmark dataset and documentation in the form of a report
D1.6 Documentation for accessing operational OS data

The ambitious objective to set-up an operational access to OS data has been achieved by different research groups in the OpenSense ecosystem, also building from the experience and networking activity that grew up during the project. However, this activity was usually conducted in the frame of research projects of feasibility studies for demonstrating the added value provided by OS data. The following is a survey of the use cases of access to OS data, based on the experience of OpenSense members. Sometimes, the data acquisition process was completed in real- or near-real time, sometimes it was offline. Recommendations for standardizing operational access to OS data is a logical next step, however, they are beyond the scope of this document. The ultimate scope of this document is to help potential OS data users and stakeholders to identify opportunities and bottlenecks in view of an operational access to OS data.

Funded by the European Union

www.cost.eu
www.opensenseaction.eu

2. Operational Access to CML data

This section reviews examples of operational access to CML data.

2.1. HoWa-PRO project (Germany)

As part of the BMBF-funded HoWa-PRO project (https://www.wasser.sachsen.de/howa-pro.html, in German), the Deutscher Wetterdienst (DWD) and Ericsson developed a real-time data pipeline for CML data for radar rainfall adjustment. Ericsson collected TSL and RSL values from several thousand CMLs at a sub-minute instantaneous sampling rate and provided the data through a cloud-based storage system in JSON format. DWD retrieved the data from this storage at one-minute intervals. To ensure over 99% data availability, each request for a given time step was delayed by two minutes. As a result, CML data were processed and ready for use 2.5-3.5 minutes after their sampling timestamp. Consequently, CML-adjusted radar fields became available approximately four minutes after the latest radar image. This acquisition pipeline operated continuously for 1.5 years, until spring 2025.

2.2. ArpaE-Lepida cooperation (Italy)

The regional agency of environmental protection of the Emilia-Romagna region (ArpaE) in Italy has an informal agreement with the local network operator Lepida to receive real-time streams of CML raw data collected from their network of around 300 links located within the region (mostly in the hilly and mountainous part of it). These data were archived until 2024. A text file with the raw data (time stamp, RLS and TSL levels and CML identifiers) was received every 15 min, with an average latency of about 7 min. Following an update of Lepida's network the process has been temporarily interrupted. However, as the network re-configuration stage has been completed and an updated metadata table was made available, ArpaE has planned to resume data recovery and set-up a processing chain for meteorological applications. Two use cases have been identified: merging of CML data and radar products (owned by ArpaE) and assimilation into models.

3. Operational Access to PWS data

This section reviews examples of operational access to PWS data.

3.1. PWS via EUMETNET (available for all member states)

EUMETNET is working at a European level to coordinate efforts on the exploitation of opportunistic data by national meteorological services. This includes the running of the E-GVAP and E-ABO programmes for GNSS water vapour and opportunistic aircraft-based observations. The E-IoT module is running a pilot project to build an alpha version of a system to share opportunistic data in near real time. This is currently sharing PWS data from Netatmo and citizen networks in France and Italy, with plans to add a network in Spain. In 2026 webcam data will be used to generate weather observations using machine learning. These will be shared via the datahub for Member organisations to use in near real time.

In the second half of 2025, EUMETNET (https://www.eumetnet.eu) began providing Personal Weather Station (PWS) data for most European countries in collaboration with Netatmo (https://weathermap.netatmo.com/). Netatmo delivers the data to MetNorway, which then distributes it via the E-SOH system, both in raw form and quality-controlled using MetNorway's titanlib.

3.2. Activity of Deutsche Wetterdienst

German national hydrometeorological service Deutsche Wetterdienst (DWD) started to collect and process PWS data within the EUMETNET program described above. During the first three months of piloting, the dataflow remained unstable and did not operate smoothly. As of September 2025, no reliable pipeline to individual meteorological services had been established, with the causes partly unknown to DWD. The data is provided in GeoJSON format and can be used at DWD for real-time radar adjustment.

3.3. Activity of The Royal Netherlands Meteorological Institute

The Royal Netherlands Meteorological Institute (KNMI) has supported research on merging of crowdsourced rain gauge and radar precipitation accumulations (Svatos, 2025), and is potentially interested in improving radar precipitation products with rain gauge data from personal weather stations. Moreover, KNMI has invested in developing quality control of PWS rain gauge data (De Vos et al., 2019) and has tested merging of PWS rain gauge and European radar data (Overeem et al., 2024). KNMI is also involved in WoW (WOW),(https://wow.knmi.nl/) and uses at least PWS and official temperature data to make temperature maps (but not precipitation).

3.4. Activity of The Swedish Meteorological and Hydrological Institute (SHMI)

At SMHI, the possibility to use PWS data for operational meteorology and hydrology has been investigated for around 5 years. Initial work focused on temperature but with time more and more attention has been put on precipitation or, more specifically, rainfall. Promising results from case studies covering different locations (cities) and (pluvial flood) events generated an interest in investigating the potential to integrate PWS in operational systems and products.

Currently (autumn 2025), work is ongoing in two separate (but related) tracks. In the national track, PWS data are integrated with observations from satellites, C-band weather radar composites and the national network of meteorological stations to produce national gridded precipitation on 2×2 km² spatial resolution and potentially sub-hourly temporal resolution. In the local/urban track, PWS data are integrated with observations from local weather radars (C- and/or X-band) and local meteorological stations to produce a higher-resolution (500×500 m², 5 min) grid with focus on rainfall for urban needs. The local/urban data will furthermore be used for short-term rainfall nowcasting. Different approaches to quality control are being developed, implemented and evaluated, including code developed within OpenSense.

Both the national and local/urban products are intended for hydrological applications. On the national scale, the data will be used as forcing (spin-up) data in the national flood forecasting model S-HYPE, with a 1-h time step. On the local/urban scale, the data will be used for urban hydrological applications including real-time control of sewer systems as well as of facilities for rainwater harvesting. Experiments are being conducted and conceivably operational production will begin in 2026.

Currently data from about 10 000 PWS are collected in real time. OpenSense has been important for the above development at SMHI, both owing to a number of STSMs and other grants, and more generally from the opportunity to network and discuss with colleagues at various meetings.

3.5. Activity of the Met Office - United Kingdom

The UK Met Office uses PWS data from the 'Weather Observations Website' (WOW) which is an initiative to collect citizen PWS observations. These data are collected, quality controlled and then used in a nowcasting system available to operational forecasters to help with issuing warnings for high impact weather such as severe convection. The data are aggregated to a time resolution of 1hour and there are approximately 2000 stations in the UK, with around 1200 used after quality control. These data are blended with data from the UK Met Office's high resolution NWP model to produce a better estimate of current conditions. There are plans to move to 15 minute resolution in the next year.

In addition to PWS data, the UK Met Office is engaged in the operational use of opportunistic GNSS data for water vapour retrieval, and aircraft-based Mode-S data. Both of these opportunistic data sources are established in operational pipelines and shared internationally, with both having large positive impacts on numerical weather prediction model forecasts.

4. Operational Access to SML data

This section reviews examples of operational access to SML data.

4.1. Nefocast project (Italy)

SML data collected in a few locations in Tuscany and in the city of Rome (Italy) are publicly accessible from the website of the Nefocast project (http://www.nefocast.it/nefocast) for the purpose of research and development. On-demand data can also be accessed via web services. The SML ground terminals were installed in the frame of the activities of Nefocast and of other research projects. Most of the above ground terminals are SmartLNB bi-directional systems operating within the the EUROBIS satellite service network managed by MBI, which currently relies on the Eutelsat

GEO satellite 8W (in the past Eutelsat 10E or Eutelsat 36E have been used as well). Ground data collected by disdrometers and rain gauges are used for SML calibration and validation purposes. The available raw data is the SNR after the DVB-S2 demodulator as well as precipitation products (1-min rainfall intensity time series) obtained from physical models (Giannetti et al., 2019) and AI/ML-based techniques (Scognamiglio et al., 2024). Additional data necessary to the conversion of the signal level into precipitation are the zero-degree isotherm height, which is provided every six hours by LaMMA, a public company providing meteorological services and owned by the Tuscany region.

4.2. HD Rain network (France)

HD Rain (https://www.hd-rain.com/) is a French company offering weather services. HD Rain deployed more than 200 SMLs in Southern France, a region, which includes densely urbanized areas, and it is prone to significant environmental and meteorological risks (Nebuloni et al., 2025). In particular, this region is exposed to “épisodes cévenols”, that is, intense Mediterranean rainfall events that can cause severe flash floods. Additionally, the complex topography of the region results in highly variable rainfall patterns, which makes conventional measurements difficult. Each HD Rain SML is equipped with a satellite dish antenna, a standard TV Low-Noise Block (LNB), a solar panel, a GSM antenna, and a device (Rainbox), which logs the RSL at a high sampling rate, and, by a SIM card, transmit data via the GSM network to a central server, where they are processed. SML terminals also host temperature and humidity sensors, which enable them to contribute to forest fire risk management. HD rain is operating SML networks for opportunistic rainfall sensing in five other countries including also mid-income countries like Ivory Coast, Georgia or Brazil.

References

[De Vos et al., 2019] L. de Vos, H. Leijnse, A. Overeem, and R. Uijlenhoet, “Quality Control for Crowdsourced Personal Weather Stations to Enable Operational Rainfall Monitoring,” Geophysical Research Letters, vol. 46, no. 15, pp. 8820–8829, 2019, doi: https://doi.org/10.1029/2019GL083731.

[Fencl et al., 2023] Fencl M, Nebuloni R, C. M. Andersson J et al., “Data formats and standards for opportunistic rainfall sensors” [version 2; peer review: 2 approved]. Open Res Europe 2024, 3:169 (https://doi.org/10.12688/openreseurope.16068.2)

[Giannetti et al., 2019] F. Giannetti, M. Moretti, R. Reggiannini, and A. Vaccaro, "The NEFOCAST system for detection and estimation of rainfall fields by the opportunistic use of broadcast satellite signals," IEEE Aerosp. Electron. Syst. Mag., vol. 34, no. 6, pp. 16–27, Jun. 2019, doi: 10.1109/MAES.2019.2916292

[MoU] Memorandum of Understanding for the implementation of the COST Action "Opportunistic precipitation sensing network" (OpenSense) CA20136, available at https://www.cost.eu/actions/CA20136/

[Overeem et al., 2024] Overeem, A., Leijnse, H., van der Schrier, G., van den Besselaar, E., Garcia-Marti, I., de Vos, L. W., 2024a. Merging with crowdsourced rain gauge data improves pan-european radar precipitation estimates. Hydrology and Earth System Sciences 28 (3), 649–668.

[Nebuloni et al.,2025] Nebuloni, R., Graf, M., Cazzaniga, G., Mercier, F., & Turko, M. (2025). The OpenSat4Weather dataset: Ku-band satellite link data for precipitation monitoring (1.0) [Data set]. Zenodo, available at https://doi.org/10.5281/zenodo.16530167

[Scognamiglio et al., 2024] Scognamiglio, G.; Rucci, A.; Vaccaro, A.; Adirosi, E.; Sapienza, F.; Giannetti, F.; Bacci, G.; Angeloni, S.; Baldini, L.; Roversi, G.; et al. Deep Learning for Opportunistic Rain Estimation via Satellite Microwave Links. Sensors 2024, 24, 6944, https://doi.org/10.3390/s24216944.

[Svatos, 2025] J.Svatos, Value of using personal weather stations in radar precipitation merging, MSc thesis Hydrology and Environmental Hydraulics Group Wageningen University, available at https://edepot.wur.nl/693405.

[WOW] "WOW Observations 2011 - 2020." Newcastle University, Dec. 19, 2022. doi: 10.25405/data.ncl.21724970.v1.