Latest developments on

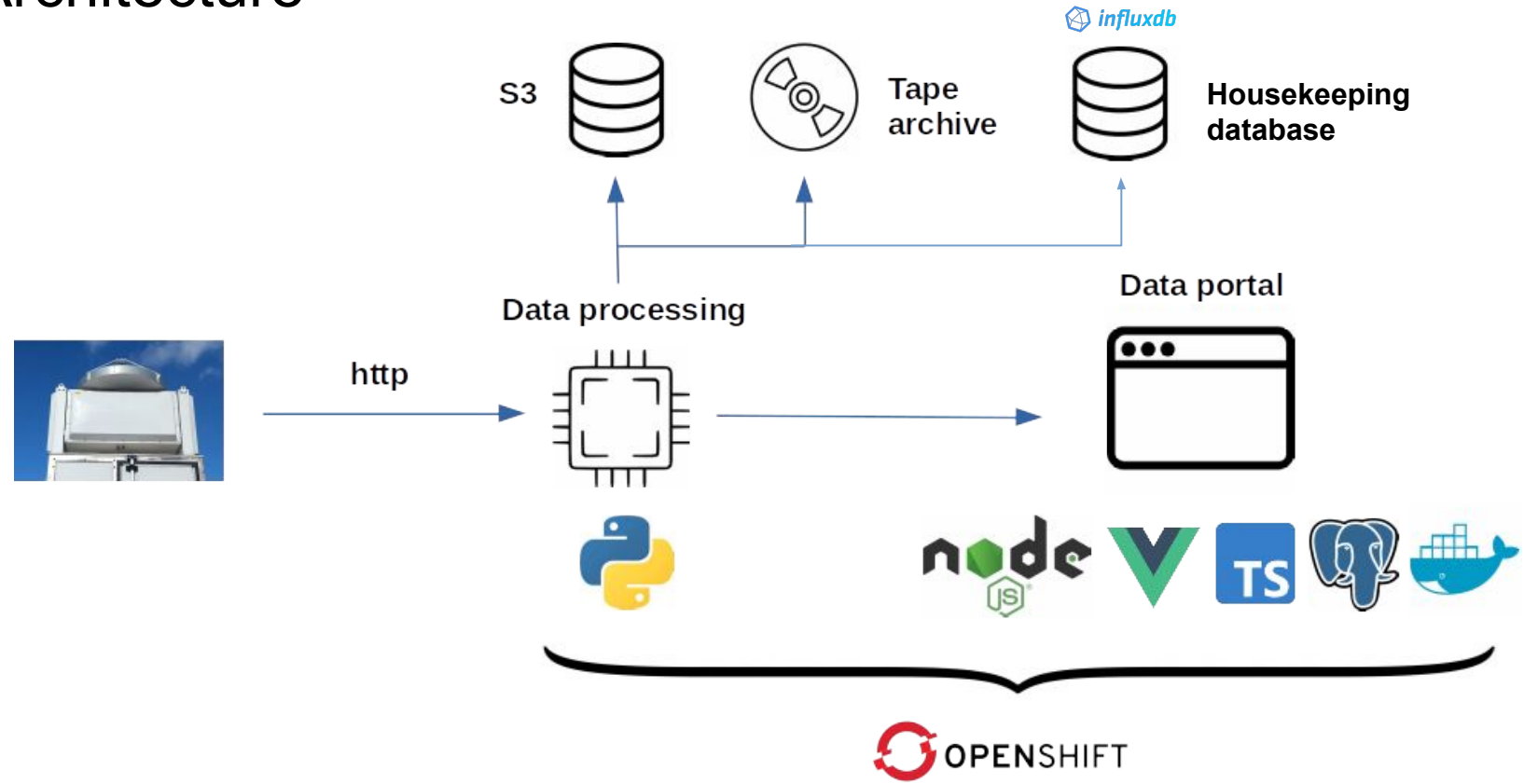
Cloudnet data portal

Simo Tukiainen, Tuomas Siipola,
Niko Leskinen, Ewan O'Connor

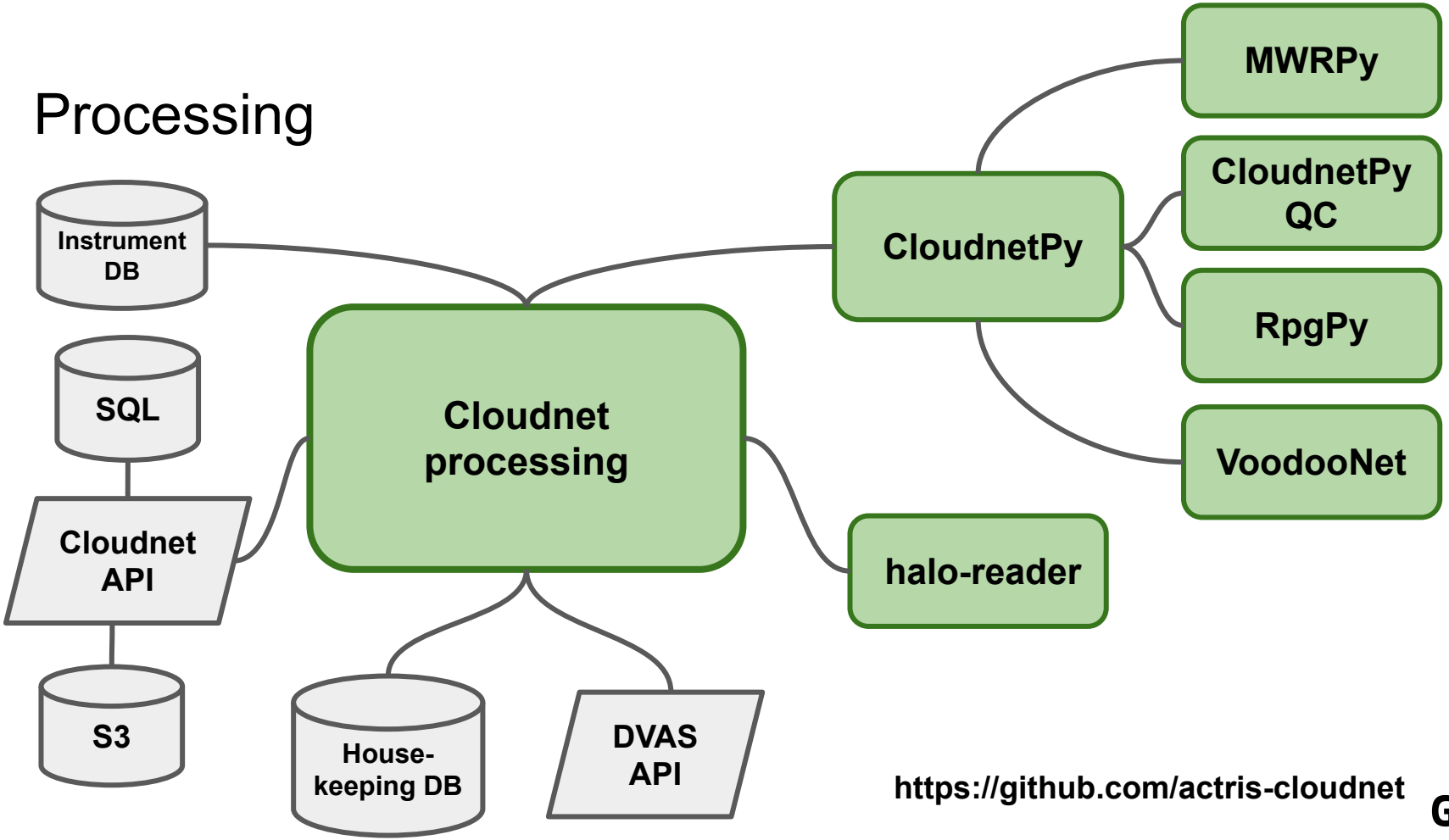
ACTRIS Data Centre – CLU unit
Finnish Meteorological Institute

CCRES workshop 2023-10-26

Architecture



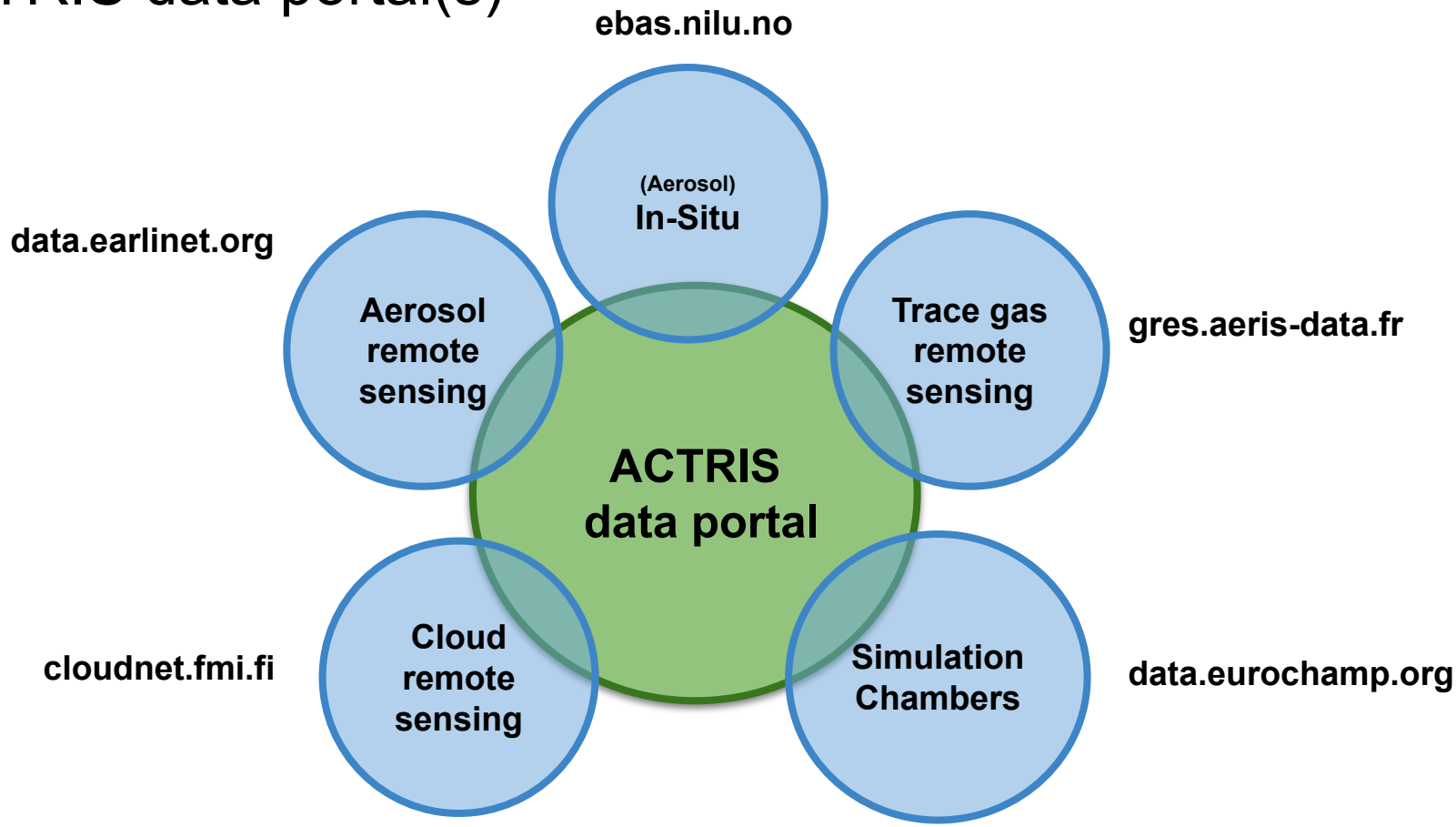
Processing



https://github.com/actris-cloudnet

GitHub

ACTRIS data portal(s)

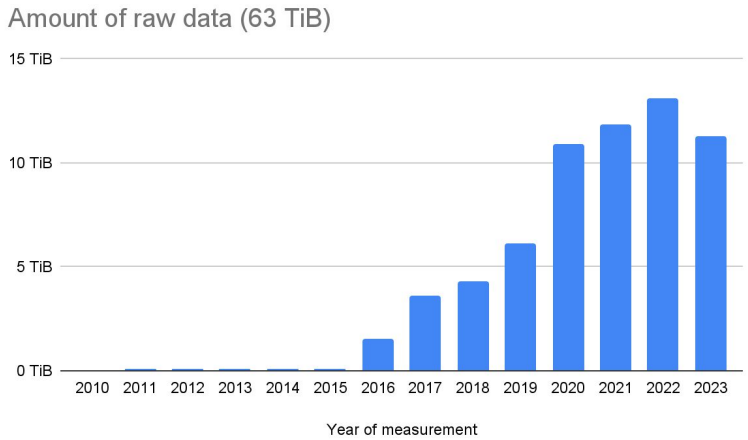


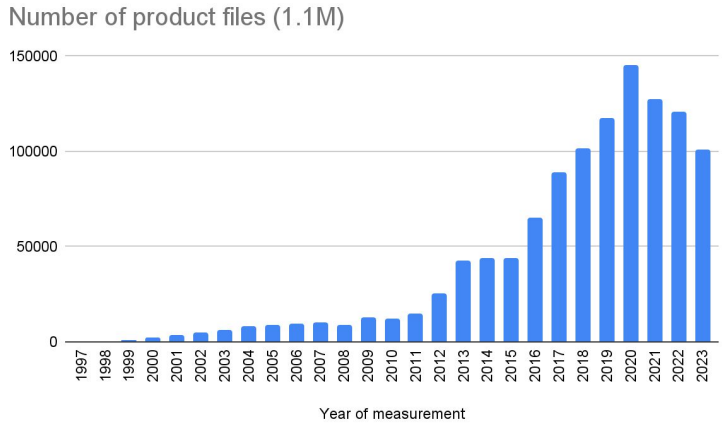
Cloudnet data volume

- ★ 971 years of data
- ★ 1.1 M product files
- ★ 6.1 M raw files
- ★ 63 TiB of raw data
  (73% RPG Level 0)

Recommended submission tool:
https://github.com/actris-cloudnet/cloudnet-submit

➔ pip install cloudnet-submit





Typical data submission problems

- Data missing (e.g. an hour at the end of date)
- Missing / incorrect instrument PID
- Invalid data

Please check logs of your submissions scripts and inspect processed data in the data portal from time to time!

Metadata

- `200 OK` : Metadata creation was successful.
- `400 Bad Request` : There was a problem in handling the request. Check request headers and content type.
- `401 Unauthorized` : Problem in authentication. Check credentials.
- `409 Conflict` : Metadata for this file already exists, and the file has been received. Do not attempt file submission.
- `422 Unprocessable Entity` : Problem in handling metadata body. Check that the metadata JSON is correct.

Data

- `201 Created` : File was received successfully.
- `200 OK` : File already exists, doing nothing.
- `400 Bad Request` : There was a problem in handling the request. Check that the MD5 checksum is valid and corresponds to the file.
- `401 Unauthorized` : Problem in authentication. Check credentials.

Cloudnet sites

Permanent sites with instrumentation required for the Cloudnet processing scheme. Most of the sites are part of the ACTRIS research infrastructure.

Site	Country	Latitude	Longitude	Altitude
Bucharest	Romania	44.348°N	26.029°E	93m
Cabauw	Netherlands	51.968°N	4.927°E	-1m
Chilbolton	United Kingdom	51.144°N	1.439°W	85m
Delft	Netherlands	51.996°N	4.379°E	-4m
Galați	Romania	45.435°N	28.037°E	40m
Granada	Spain	37.164°N	3.605°W	680m
Hyytiälä	Finland	61.844°N	24.288°E	174m
Jülich	Germany	50.908°N	6.413°E	111m
Kenttärova	Finland	67.987°N	24.243°E	345m
Leipzig	Germany	51.353°N	12.435°E	126m
Leipzig LIM	Germany	51.333°N	12.389°E	126m
Lindenberg	Germany	52.208°N	14.118°E	104m
Lutjewad	Netherlands	53.24°N	6.21°E	1m
Mace Head	Ireland	53.326°N	9.9°W	16m
Mindelo	Cabo Verde	16.8778°N	24.995°W	13m
Munich	Germany	48.148°N	11.573°E	538m
Norunda	Sweden	60.086°N	17.479°E	46m
Ny-Ålesund	Norway (Svalbard)	78.923°N	11.922°E	19m
Palaiseau	France	48.716°N	2.212°E	156m
Payerne	Switzerland	46.813°N	6.944°E	491m
Potenza	Italy	40.601°N	15.724°E	760m
Rzecin	Poland	52.758°N	16.31°E	57m
Schneefernerhaus	Germany	47.417°N	10.977°E	2653m
Warsaw	Poland	52.21°N	20.98°E	112m

Operational site · Some data · Inactive

https://cloudnet.fmi.fi/sites

Instrument PIDs

Fully operational

- Used in provenance, calibration, housekeeping and data citation
- Now required in data submission
- 164 instruments in total

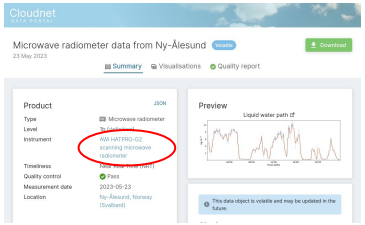
Check that information is up to date for your instruments!

- Send us photographs!

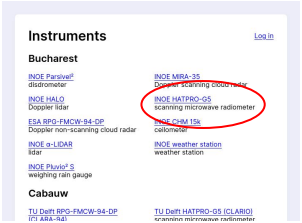
Ideas for further development?

- Login for PIs to update information themselves?
- Instrument logbook?

cloudnet.fmi.fi



instrumentdb.out.ocp.fi



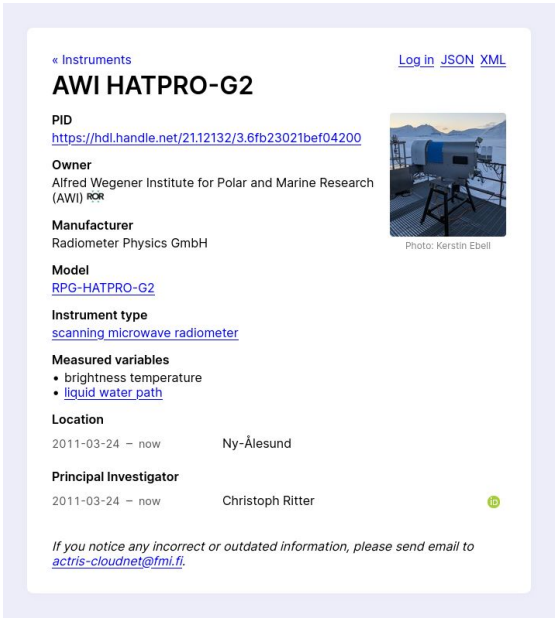
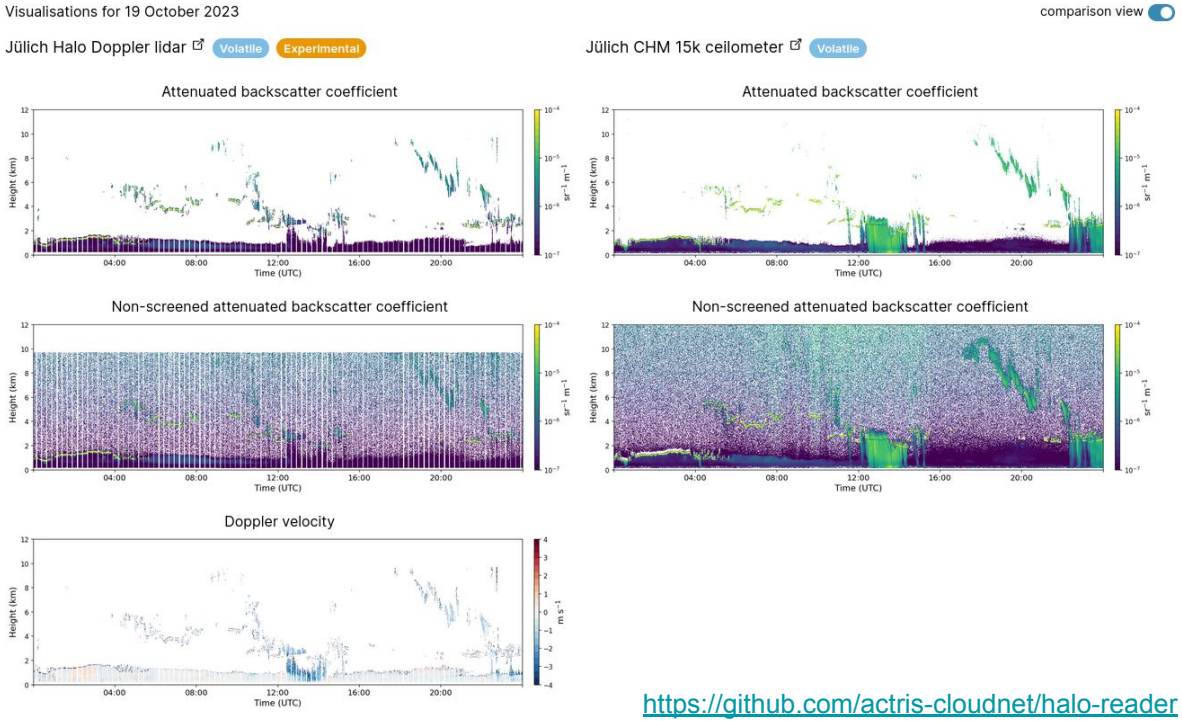


Photo: Kerstin Ebell

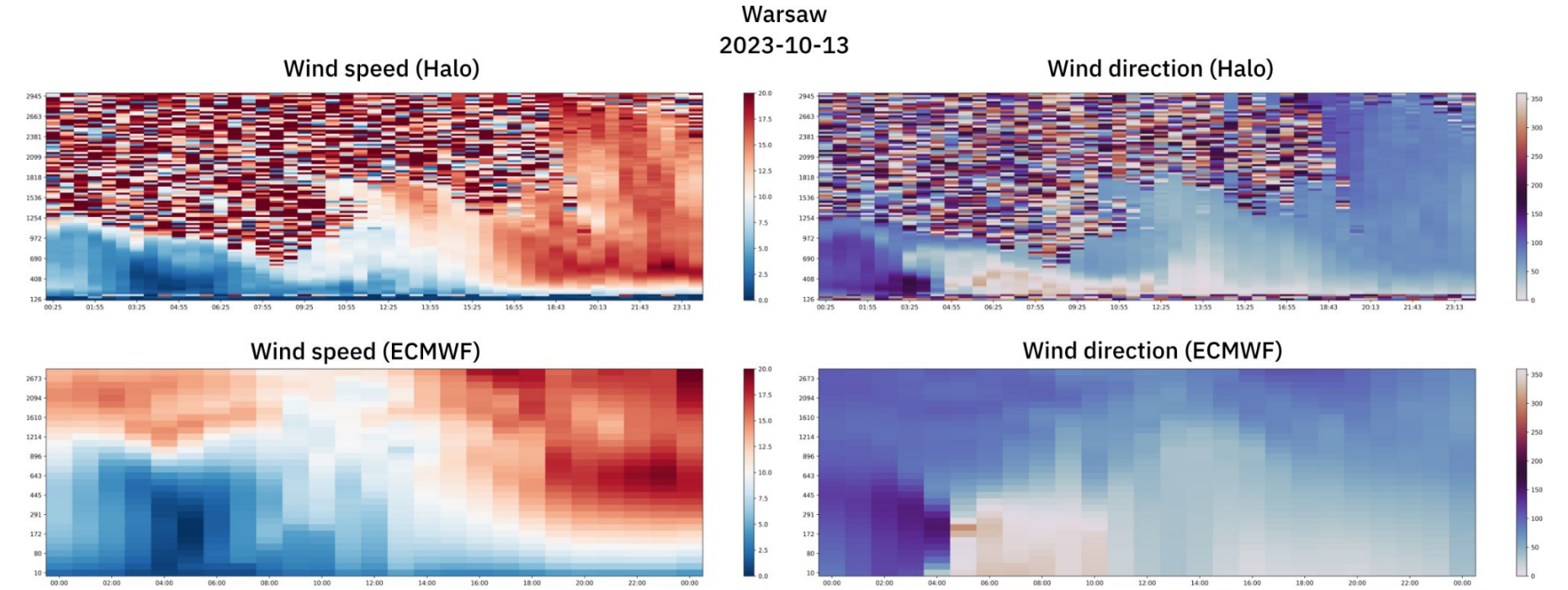
Doppler lidar processing

Vertical backscatter and velocity profiles from HALO Photonics StreamLine Doppler lidars:

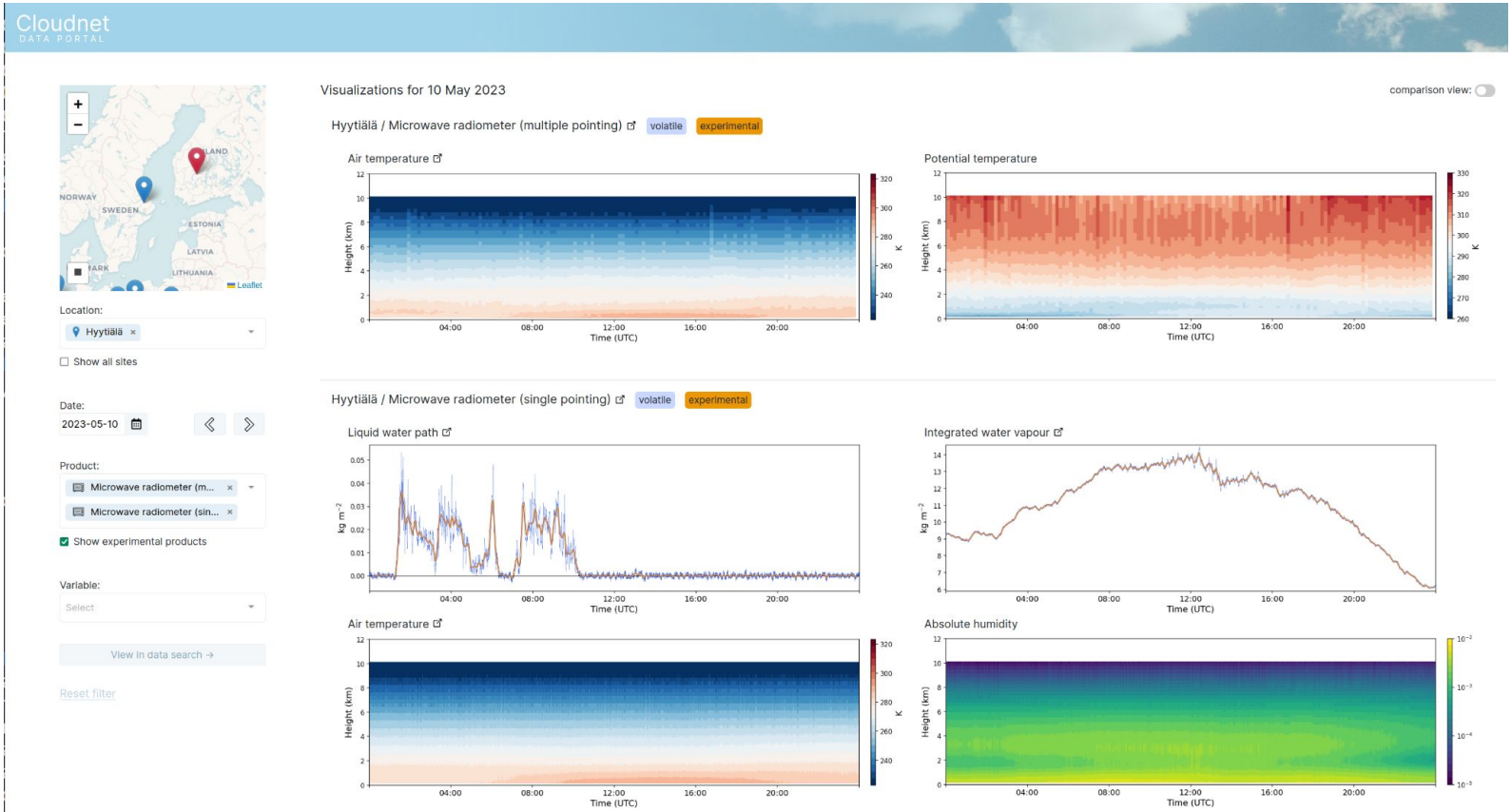


https://github.com/actris-cloudnet/halo-reader

Doppler lidar processing: Winds derived from scans



MWRpy – New processing pipeline for HATPRO



MWRPy input

- Retrieval coefficients
  - Send to Simo / Tobias

- HATPRO binary files
  - .LWP, .IWV, .HKD, .BRT, .MET, .BLB / .BLS, .IRT
  - Submit these!

Cloudnet DATA PORTAL | Search data | Visualise data

Location

Hyytiälä ×

Show all sites

Date

2023-10-19

Product

Select

Show experimental products

Instrument

Select

Variable

Liquid water path ×

Liquid water path ×

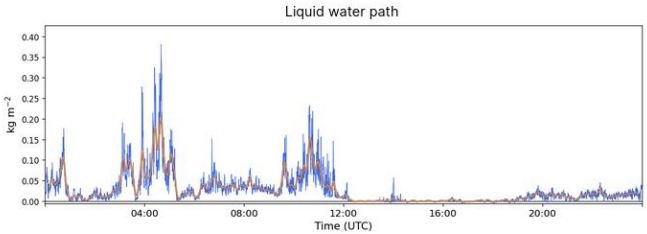
Liquid water path ×

View in data search →

Reset filter

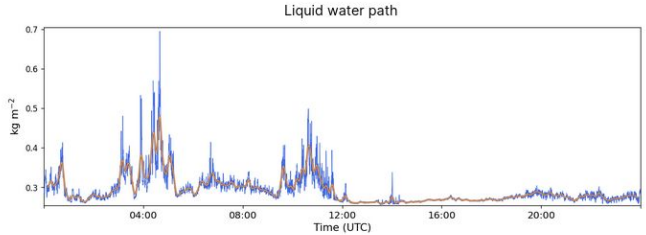
Visualisations for 19 October 2023

Hyytiälä HATPRO microwave radiometer Volatile



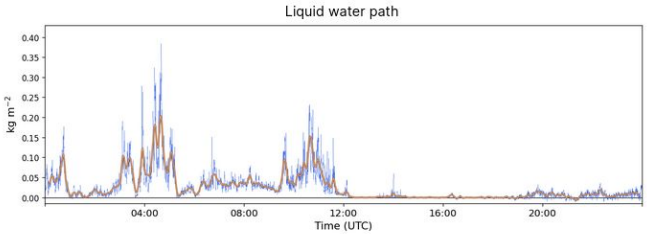
LWP from HATPRO

Hyytiälä RPG-FMCW-94 cloud radar Volatile



LWP from FMCW-94 cloud radar

Hyytiälä MWR single pointing Volatile Experimental

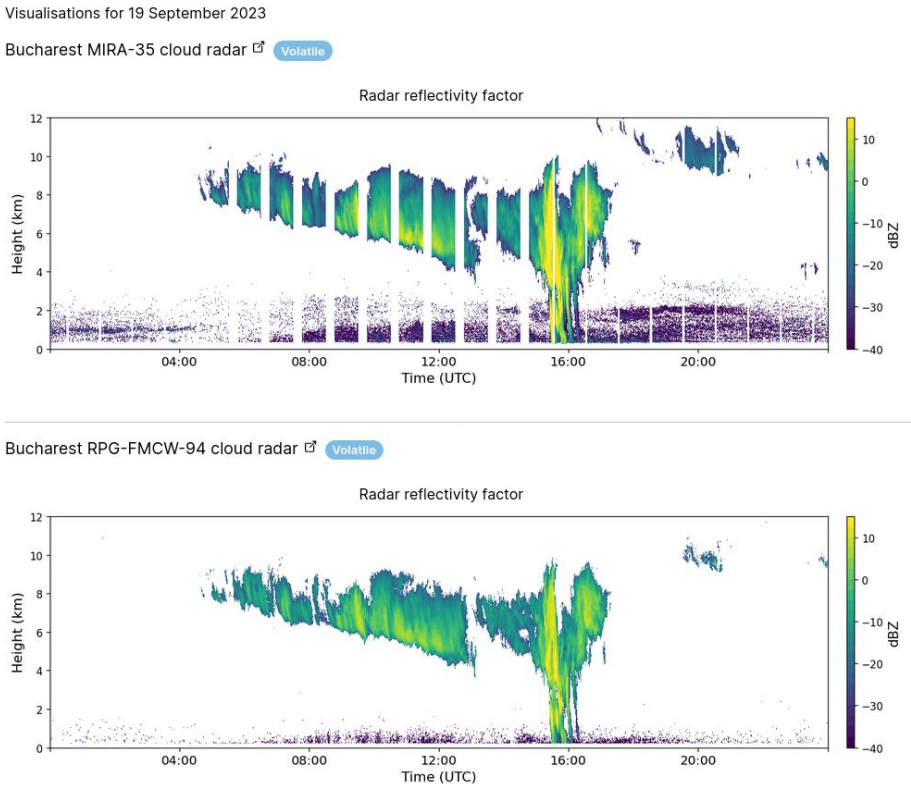


Calibrated LWP from HATPRO using MWRpy

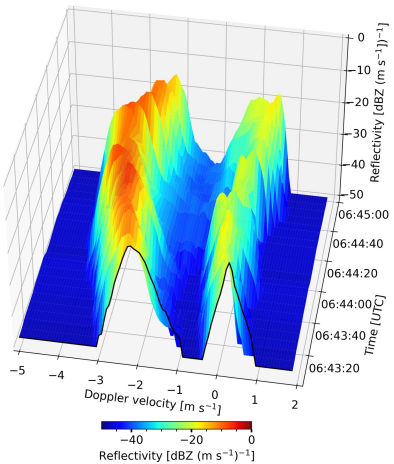
Multiple instruments

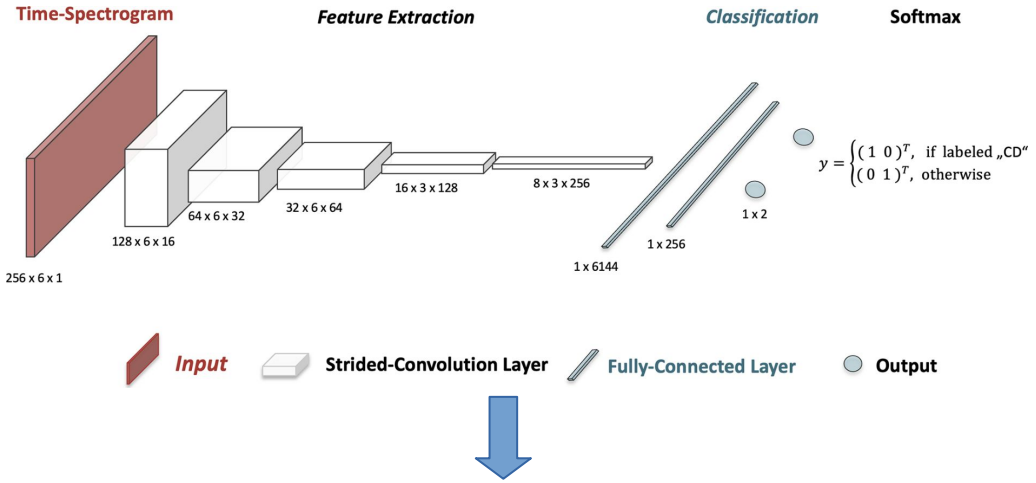
Now multiple instruments of the same type are supported for a single site (e.g. two cloud radars).

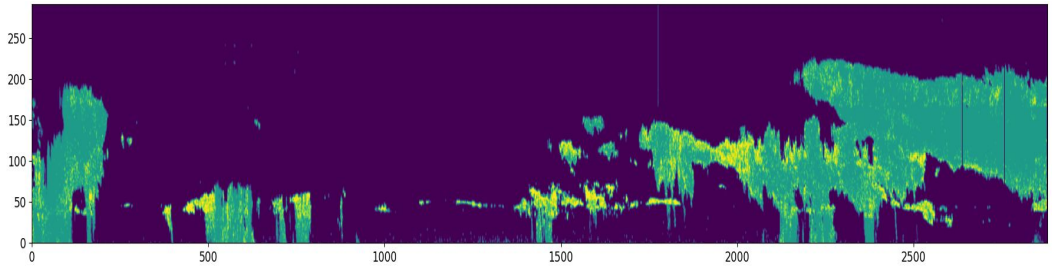
Only a single instrument is used in further processing. How to prioritise the different instruments and instrument types??



Voodoo – ML method for finding mixed-phase clouds



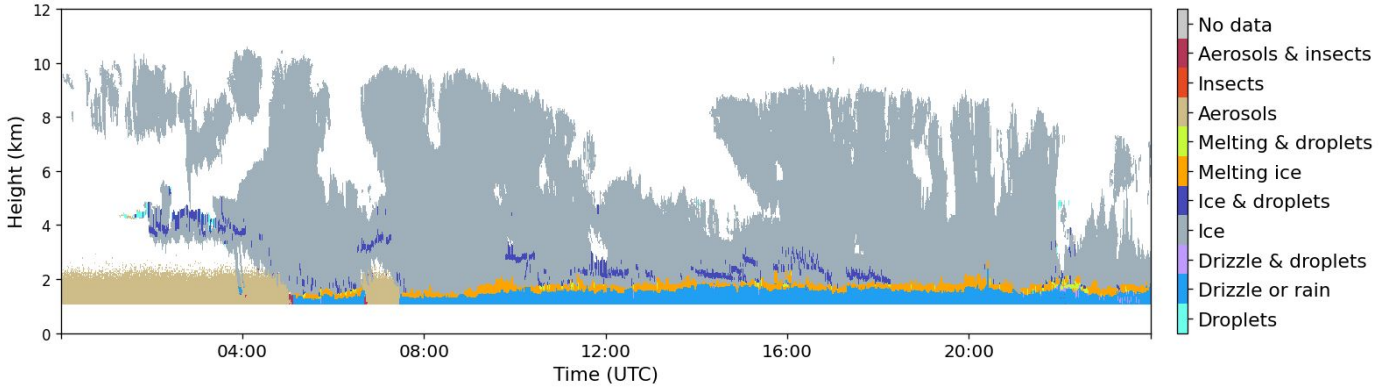




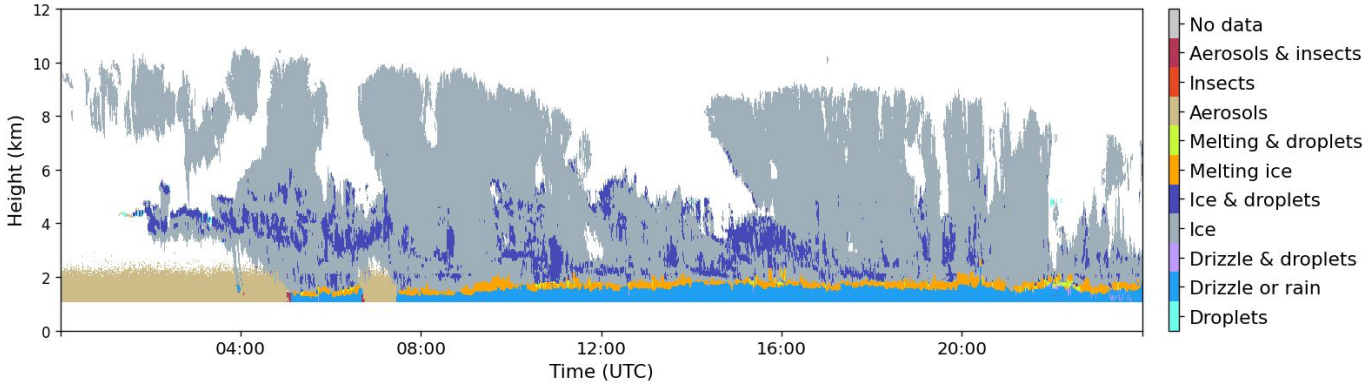
Liquid probability

Schimmel et al. (2022). Identifying cloud droplets beyond lidar attenuation from vertically pointing cloud radar observations using artificial neural networks. *Atmos. Meas. Tech.*, *15*(18), 5343-5366

Cloudnet classification



Voodoo classification



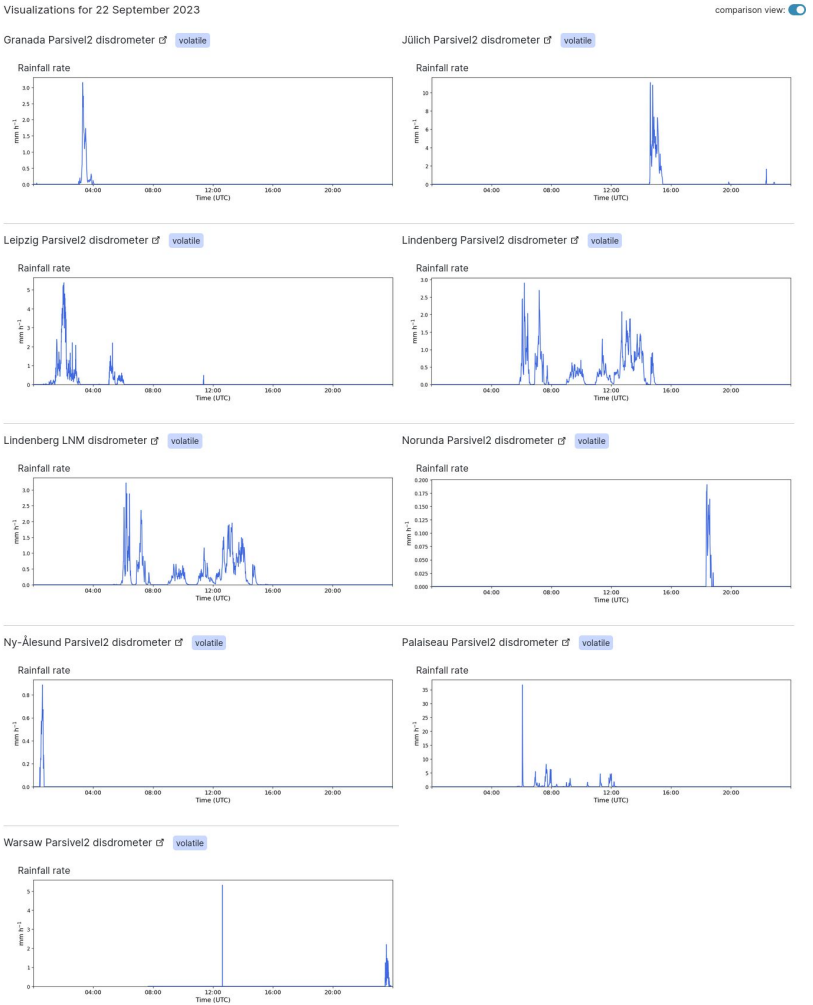
Disdrometer processing

NRT data from 8-9 sites

Parsivel² has different file format for almost all sites:

- ➔ Unspecified columns / data telegram
- ➔ Files with no timestamps
- ➔ Configuration changes without notice

We can manage for now but an effort to harmonise data logging and submission would simplify our work



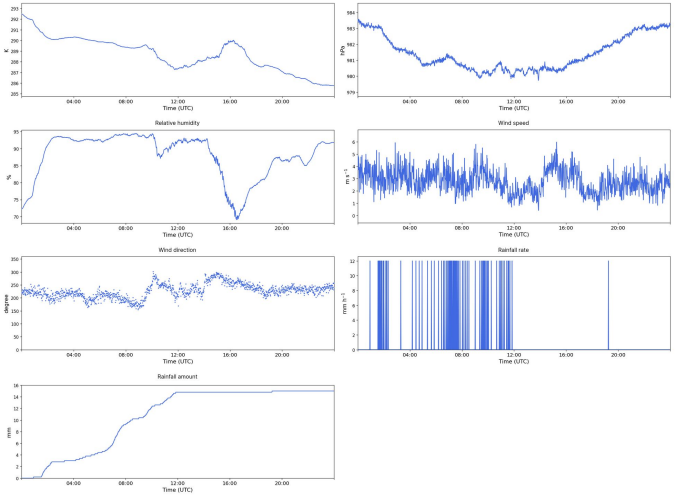
Weather station

Currently we have some data in Palaiseau/SIRTA format.

We propose a common format for all sites:
https://github.com/actris-cloudnet/cloudnetpy/blob/main/specs/weather-station.md

Requires conversion of your data before submission to the Cloudnet data portal.

Palaiseau, 21 September 2023



Collection DOIs

- DOI is now assigned for collection of files (created when clicking “Download all” on search page)
- Simplifies data citation for researchers and allows better tracking of data usage

https://doi.org/10.60656/aca60d45fe564093

Data availability

The data used in this study are generated by the Aerosol, Clouds and Trace Gases Research Infrastructure (ACTRIS) and are available from the ACTRIS Data Centre using the following link: https://doi.org/10.60656/aca60d45fe564093.

Acknowledgements

We acknowledge ACTRIS and Finnish Meteorological Institute for providing the data set which is available for download from https://cloudnet.fmi.fi. Measurements were supported by the European Space Agency through the FRM4RADAR project (ESA Contract No. 4000122916/17/I-EF). The measurements at Norunda were also supported by ICOS Sweden. We acknowledge ECMWF for providing IFS model data.

Citation

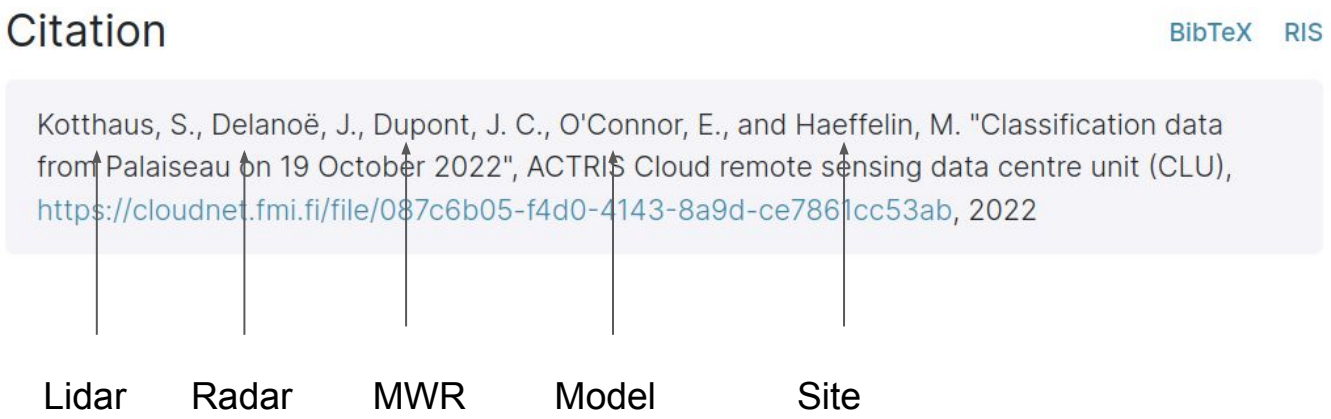
Adam, M., Constantin, D., Delanoë, J., Dupont, J.-C., Eliasson, S., Kotthaus, S., Moisseev, D., Mölder, M., Navas Guzmán, F., O'Connor, E., Pfitzenmaier, L., Pîrloagă, R., Pospichal, B., Rosu, A., Schween, J., Seifert, P., Țoancă, F., Unal, C., Walden, C., Zinner, T., Alados-Arboledas, L., Fomba, K., Haeffelin, M., Hyvärinen, A., Löhnert, U., Mayer, B., Petäjä, T., & Vasilescu, J. (2023). Custom collection of categorize, classification, disdrometer, drizzle, ice water content, and 5 other products from Bucharest, Chilbolton, Galați, Granada, Hyytiälä, and 11 other sites on 9 Aug 2023. ACTRIS Cloud remote sensing data centre unit (CLU). https://doi.org/10.60656/aca60d45fe564093

Citation

Currently the following people are included:

- Instrument PI(s)
- ACTRIS NF PI (from labelling)
- Additional site-specific people

IS THIS SUFFICIENT??

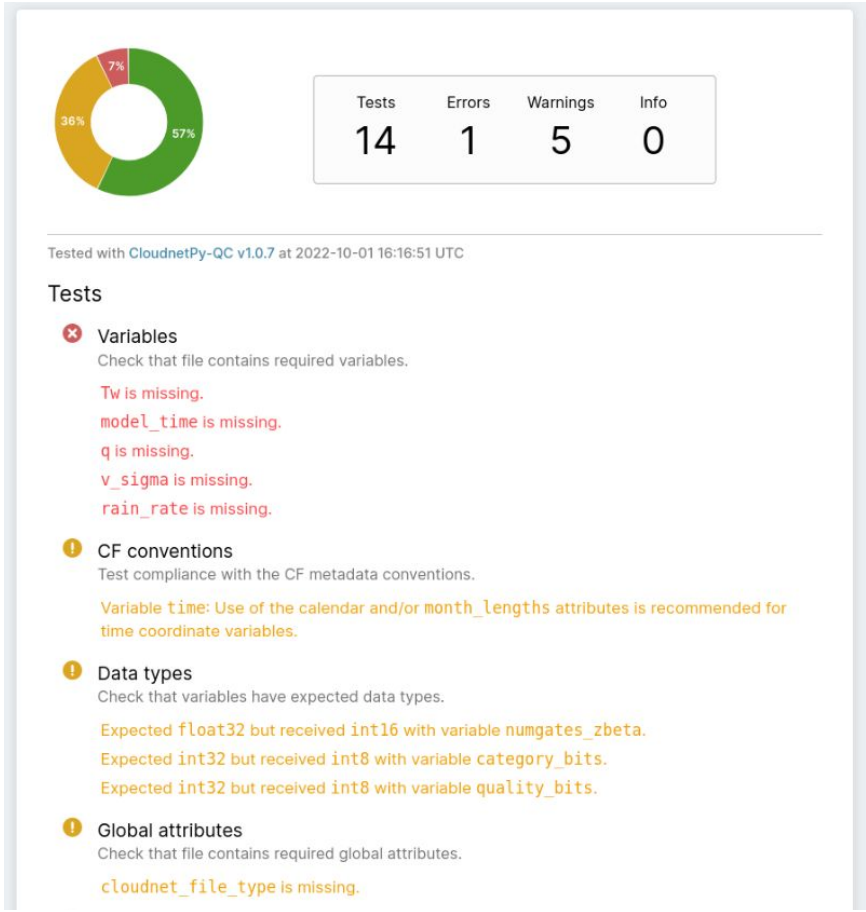


Quality control

https://github.com/actris-cloudnet/cloudnetpy-qc

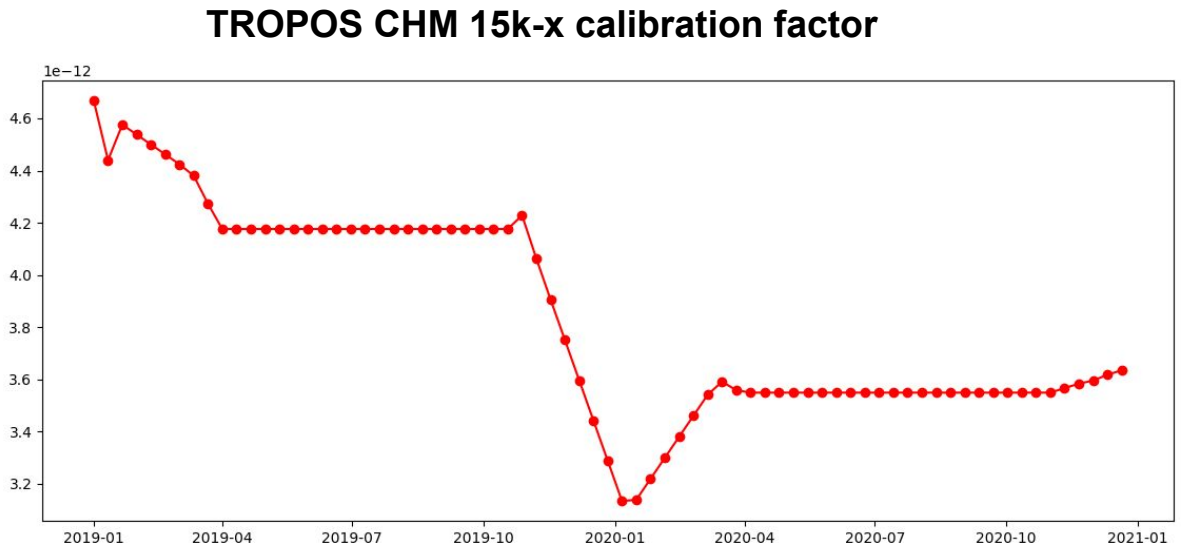
Evaluates the overall quality of a processed file:

- Metadata
- Data checks



Calibration: current

Database and API (GET, POST) for calibration data.



Calibration: next steps

- Build better architecture for calibration
- Discuss with domain experts and port their research code to operational use
- Start with ALC (Vaisala and Lufft ceilometers) and HATPRO with MWRpy

Future

- New Doppler lidar (HALO, WindCube) products
  - Winds, turbulence, …
- Utilize new data:
  - MWRPy products
  - Disdrometer data
  - Weather station data
- (More) automated Level 2 product generation
- Proper calibration workflow
- Monitoring pages for raw data submission / processing status
- Dynamic plots
- Spectral processing
- AI / machine learning