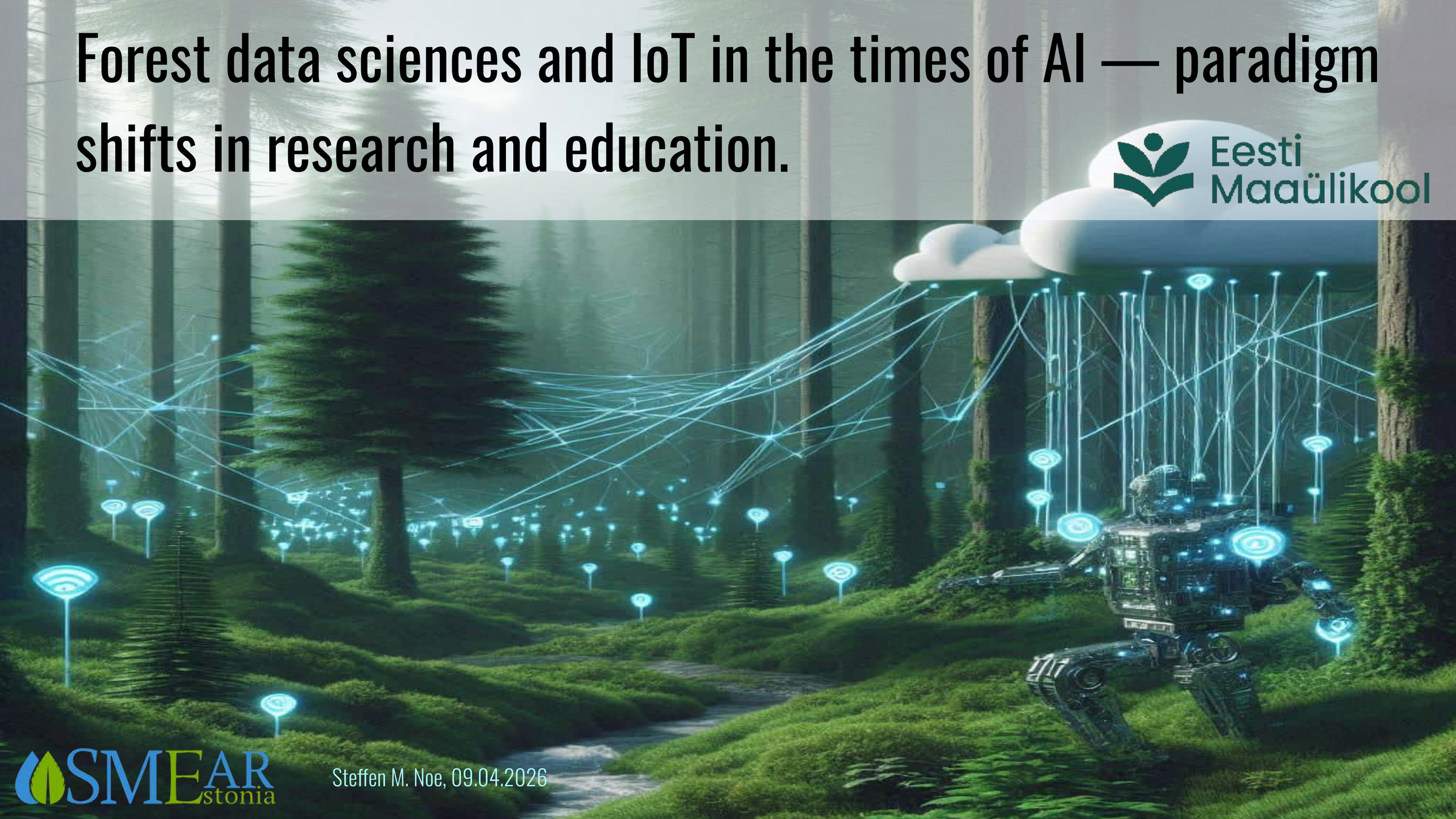
Forest data sciences and IoT in the times of AI — paradigm shifts in research and education.



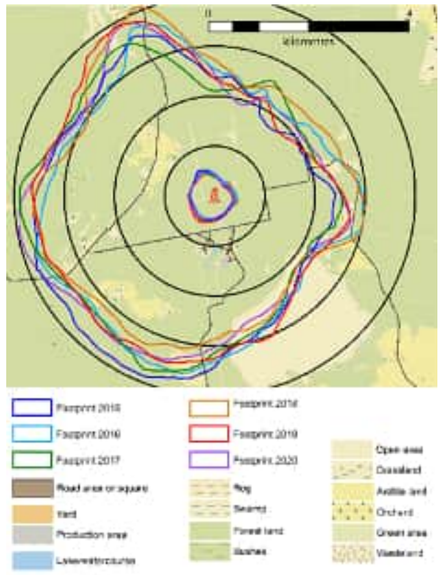
Eesti Maaülikool

SMEAR Estonia

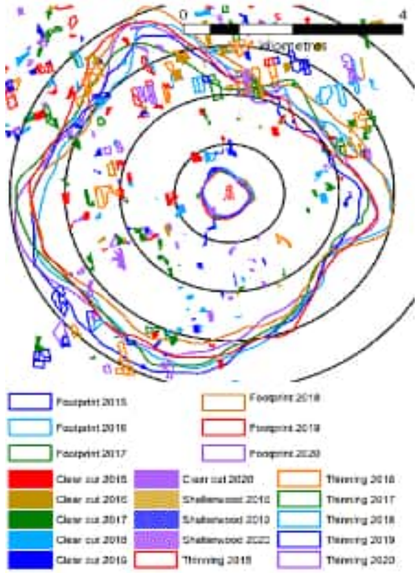
Steffen M. Noe, 09.04.2026

Spatially explicit time-series of changes in forest ecosystems

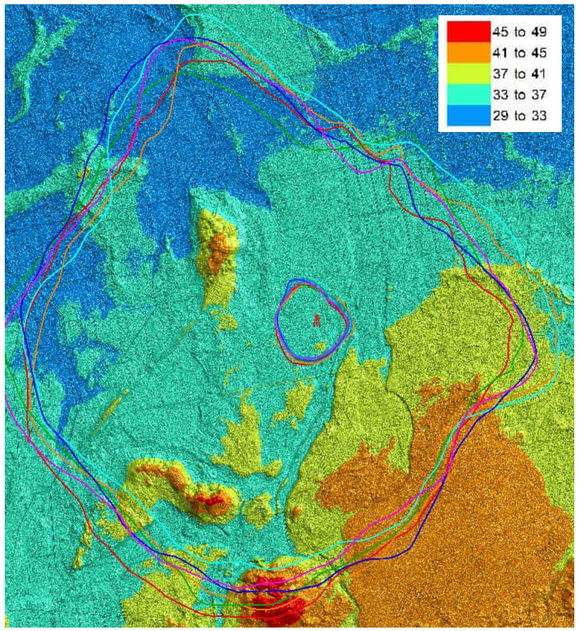
Footprint area: natural change ~1.5-5%, human induced change ~2-2.5% per year



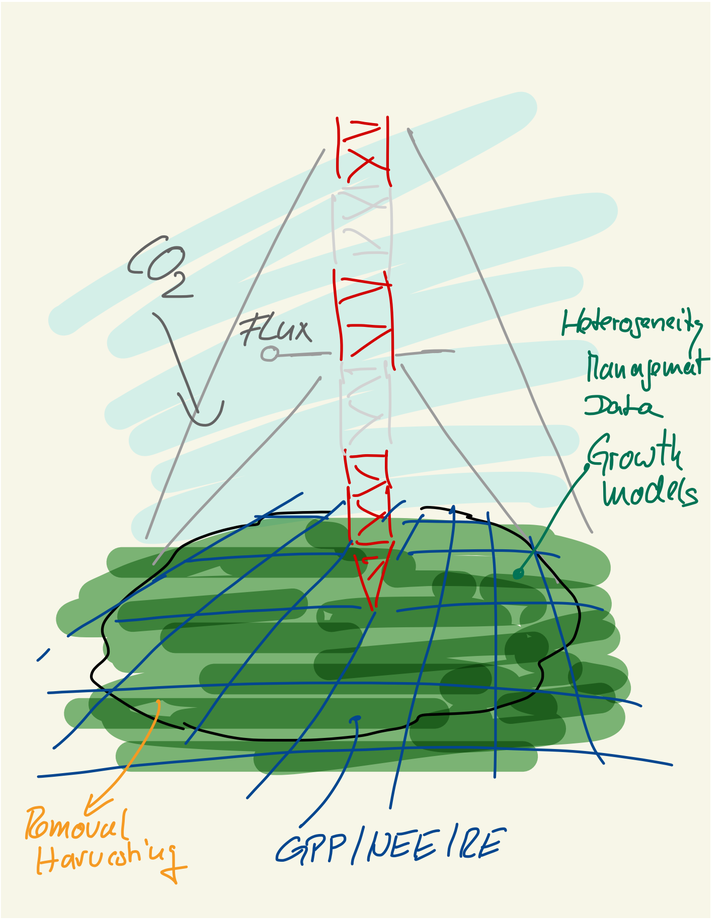
Six years of changes in the 70 and 30 meter footprint of the SMEAR Estonia station determined by wind speed and wind direction.



Six years of changes in the 70 and 30 meter footprint of the SMEAR Estonia station determined by forest management activities



Tree heigh map of the footprint area given by airborne Lidar data. These can be used to modulate the footprint calculation and to verify modelled changes in height growth in a 4 year interval.

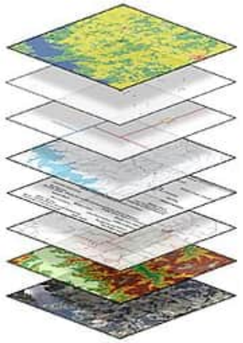


Joonas Kollo, Allar Padari

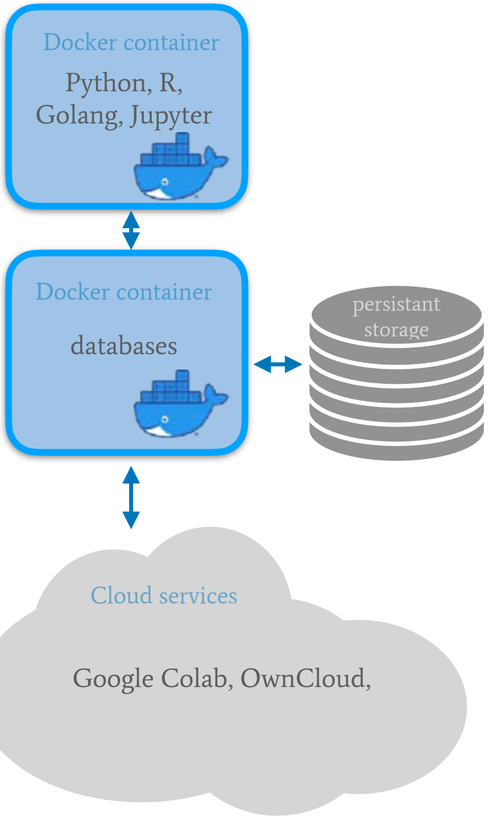
Workflow from multiple data sources towards gridded results

Benefits, the neural network can be easily automatized as microservice, no manual parameter estimations

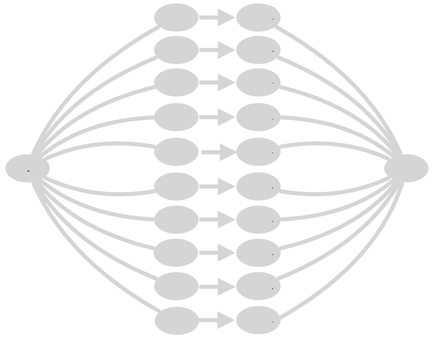
Eddy covariance, meteo, atmospheric data

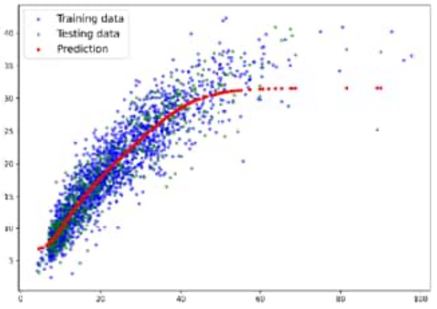


GIS data, land cover, forest management



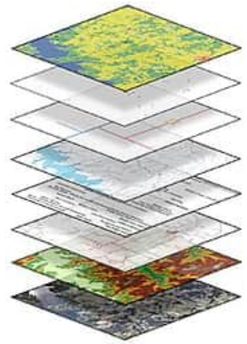
Compiling the data





sklearn, PyTorch

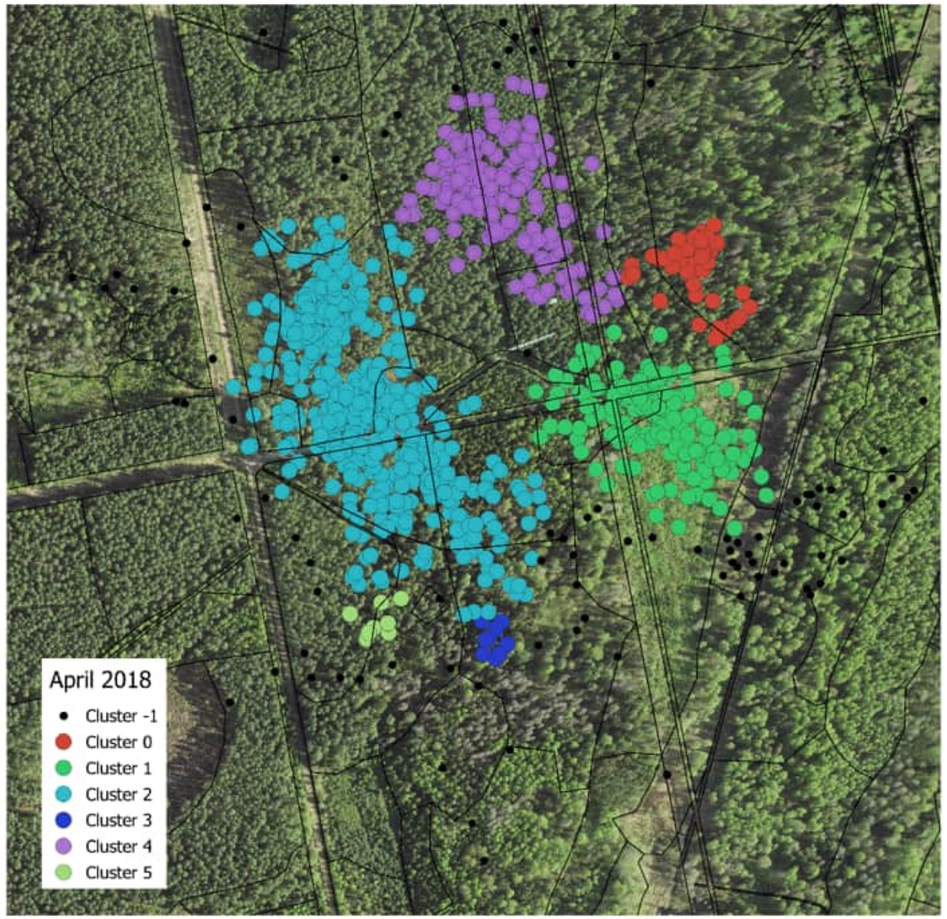
ML/AI layer

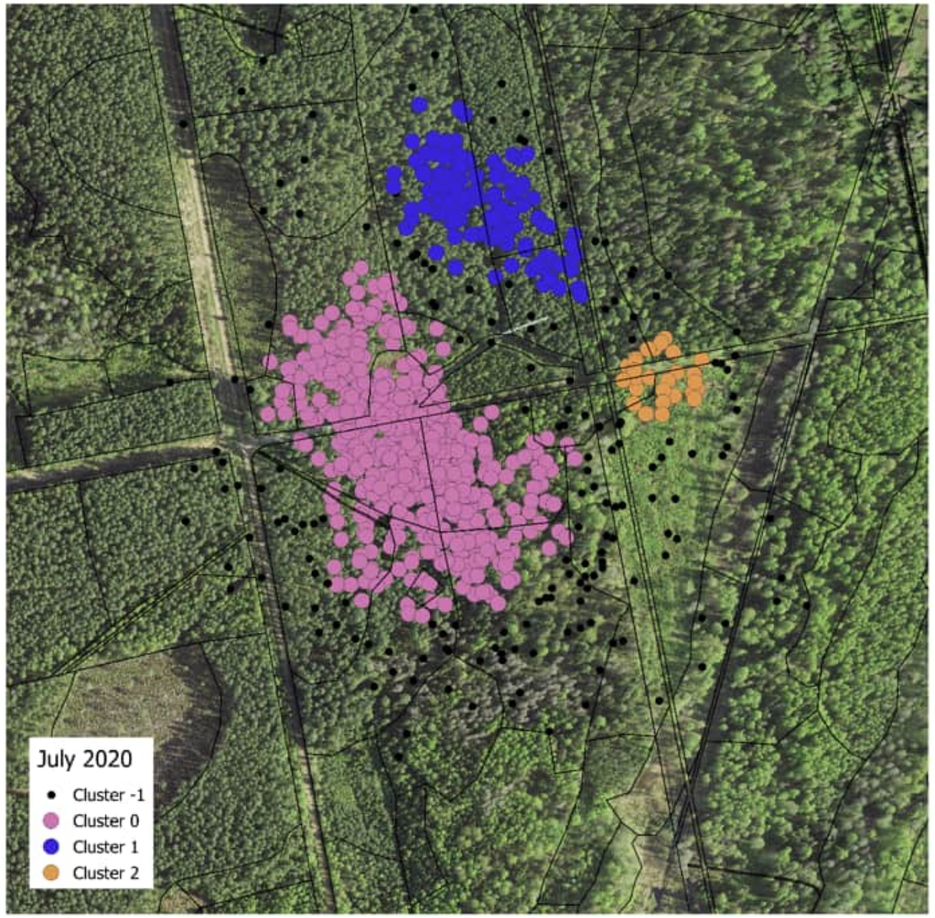


Mapping to the grid

Machine learning cluster detection of area of max. CO2 flux in the footprint

Hierarchical Density-based spatial clustering of applications with noise (HDBSCAN)





Anuj Thapa Magar, Allar Padari, Alisa Krasnova

Sensors and IoT in Forestry

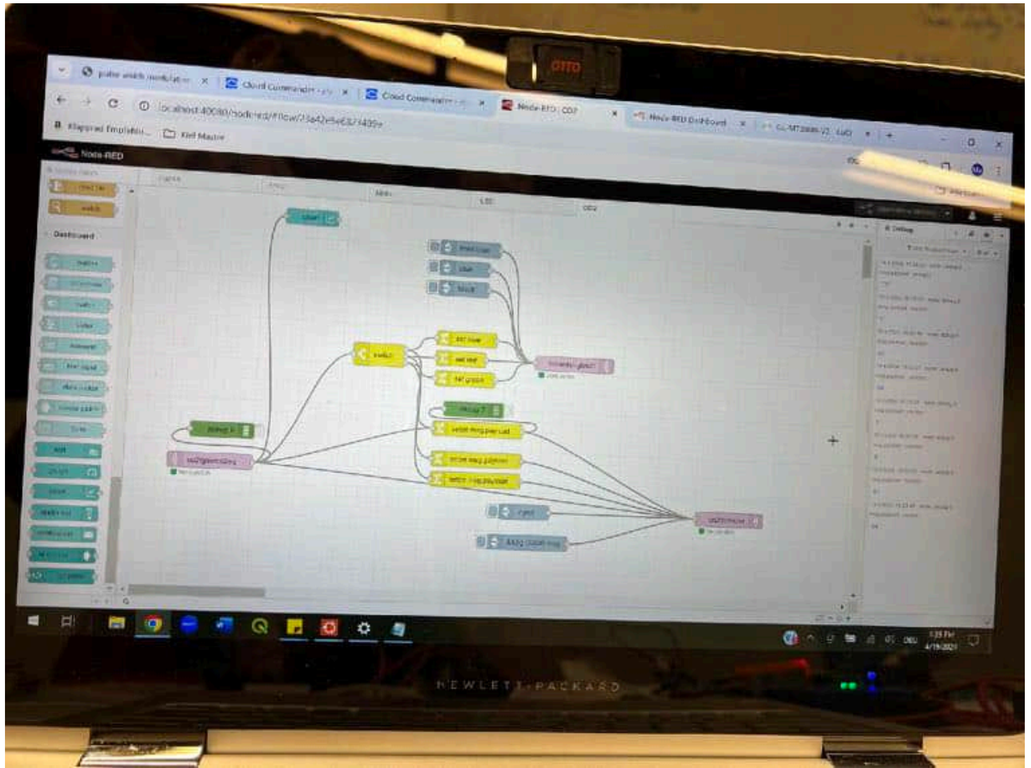
Developing low cost sensor systems that can be deployed over large areas

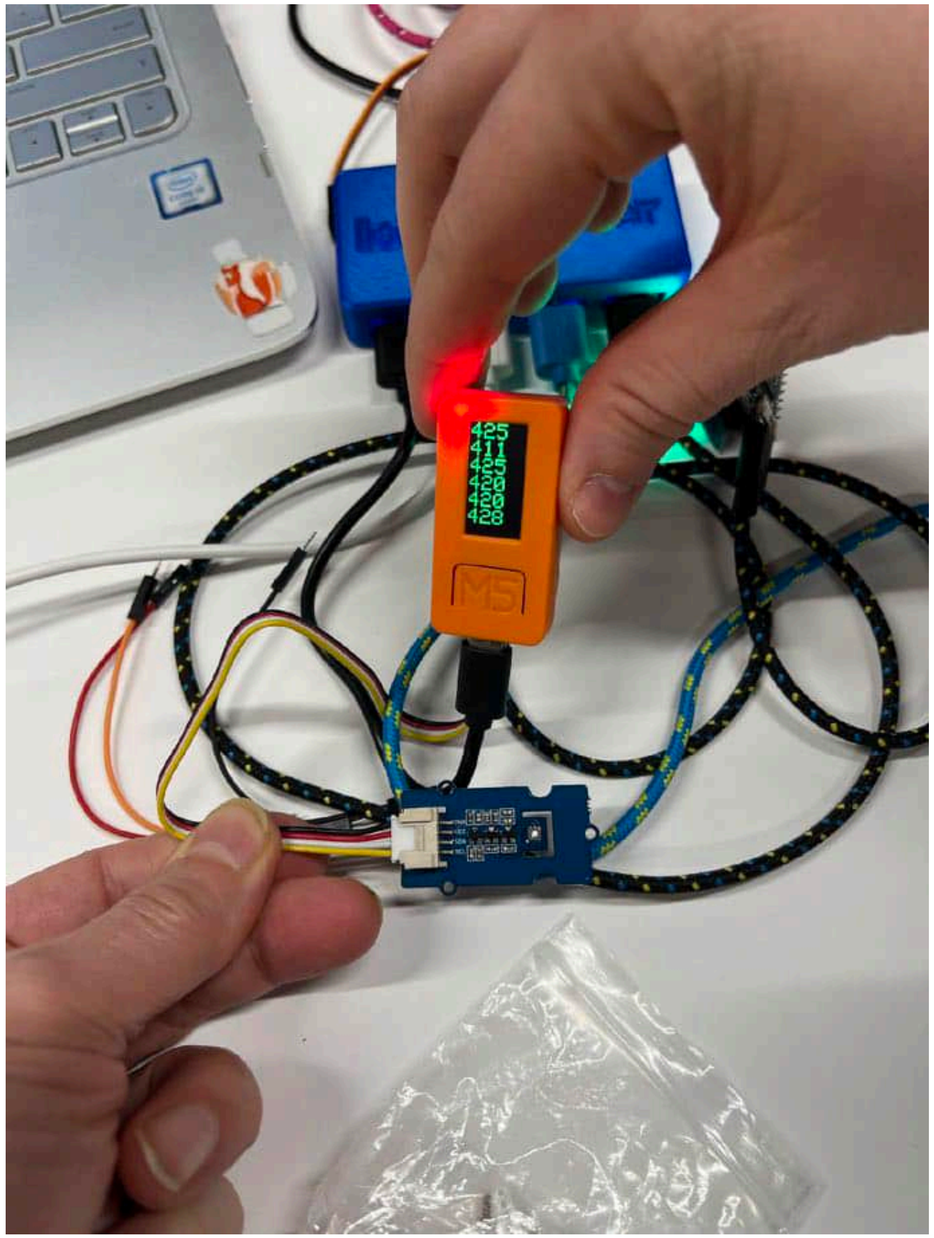
Example: Low cost CO2 sensor using the IoTempower framework

SCD4x CO2 sensor, IoTempower framework, Node Red graphical programming environment.

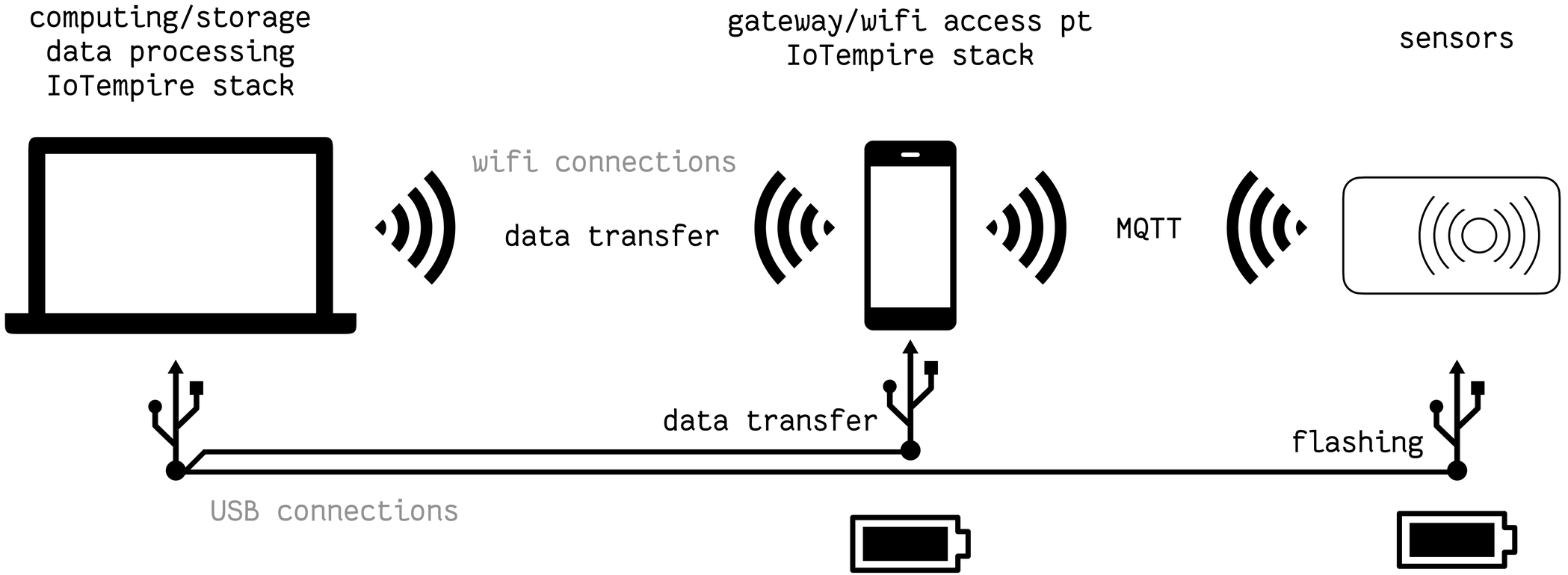
https://github.com/iotempire/iotempower/

Cooperation with Distributed Systems Group Tartu University, Prof. Ulrich Norbisrath, MSc Renato Perotto, MSc Fedir Kyrychenko





Principle setup for a CO2 sensor protoype

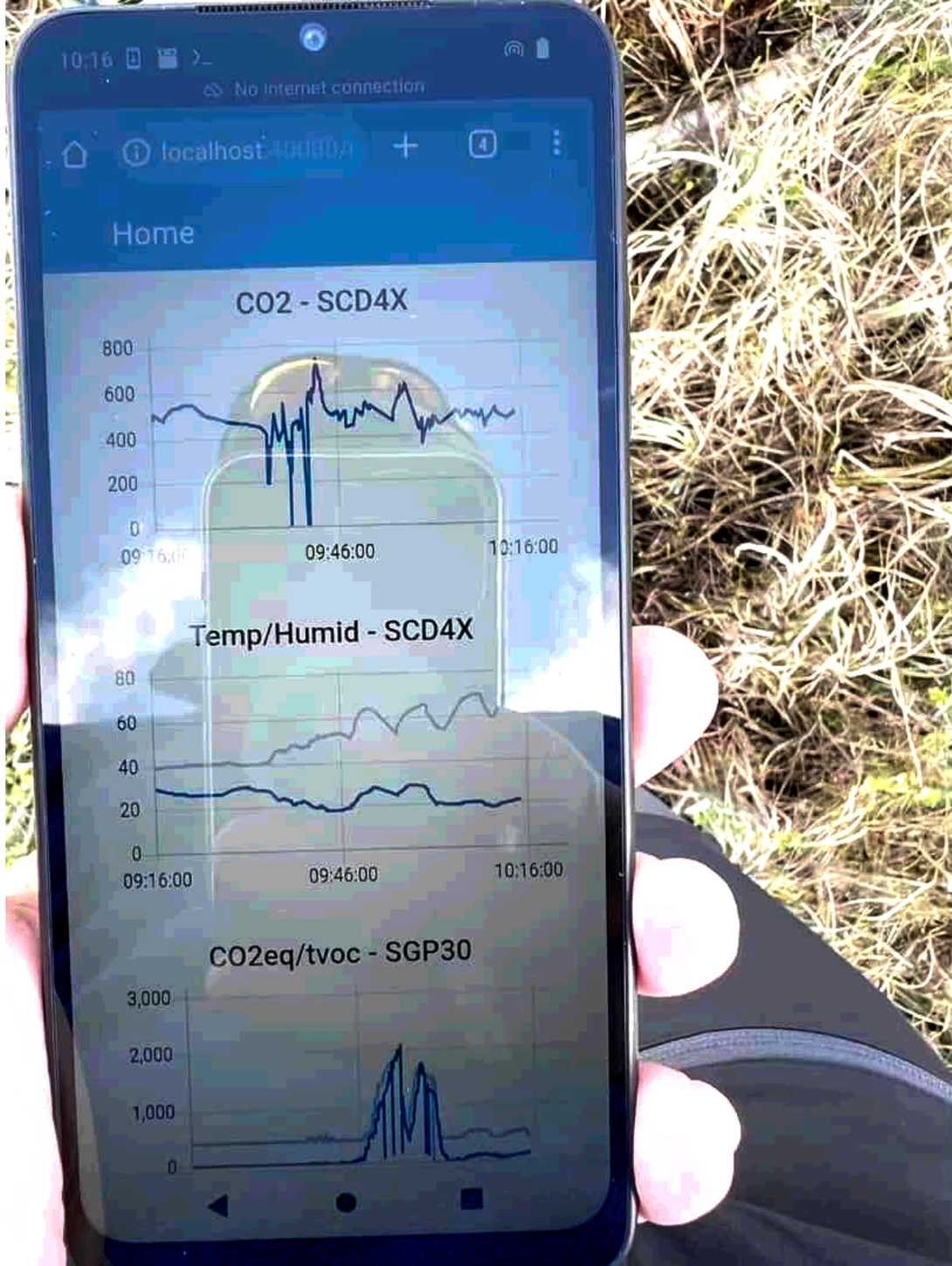


Mobile phone as gateway

https://github.com/iotempire/iotempower

- Cheap (~70-120 Euro) Android phone
- running Termux to allow Linux system behaviour
- Running the IoTempower stack
- Graphical feedback with screen for online tracking of measurements
- Graphical programming using Node Red
- local storage for short-term measurements
- data transfer to storage/server when connected to network (wifi/GSM/USB)

https://termux.dev/en/

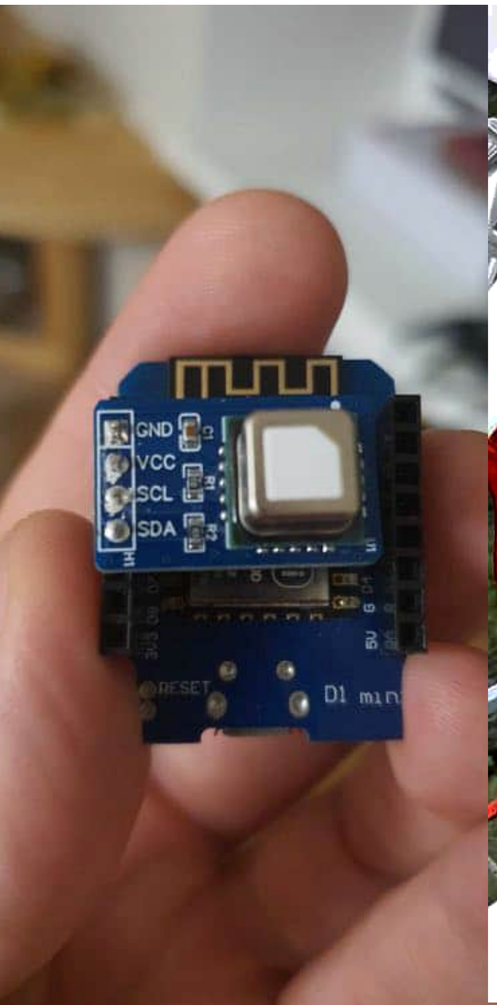


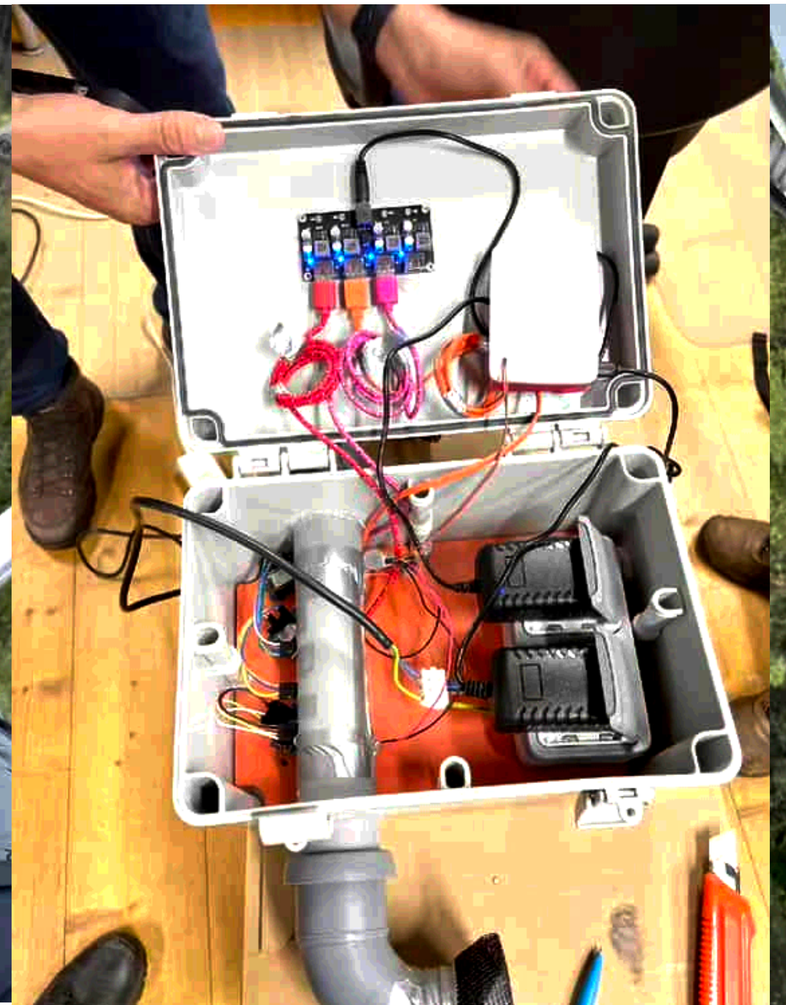
IoT sensors prototypes

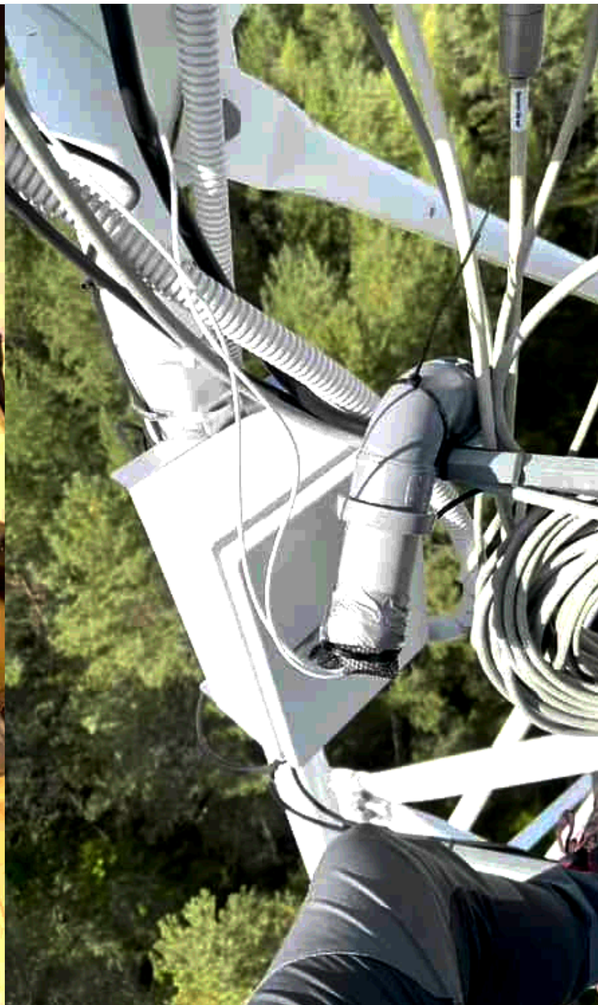
Tested during field experiment in Oulanka National Park

Testing CO2 sensors at SMEAR Estonia

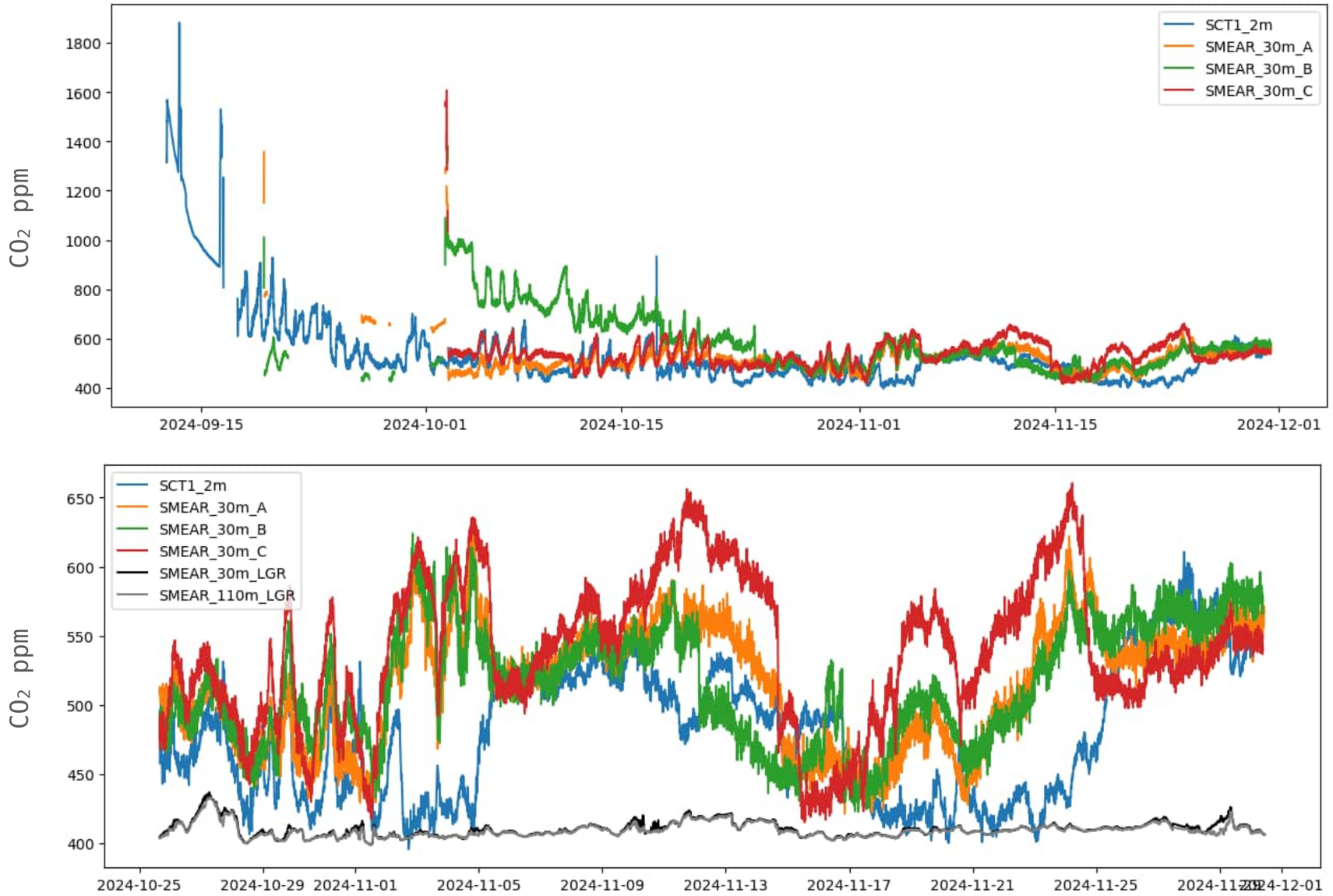
Co-location calibration by placing low-cost and high accuracy systems at the same place





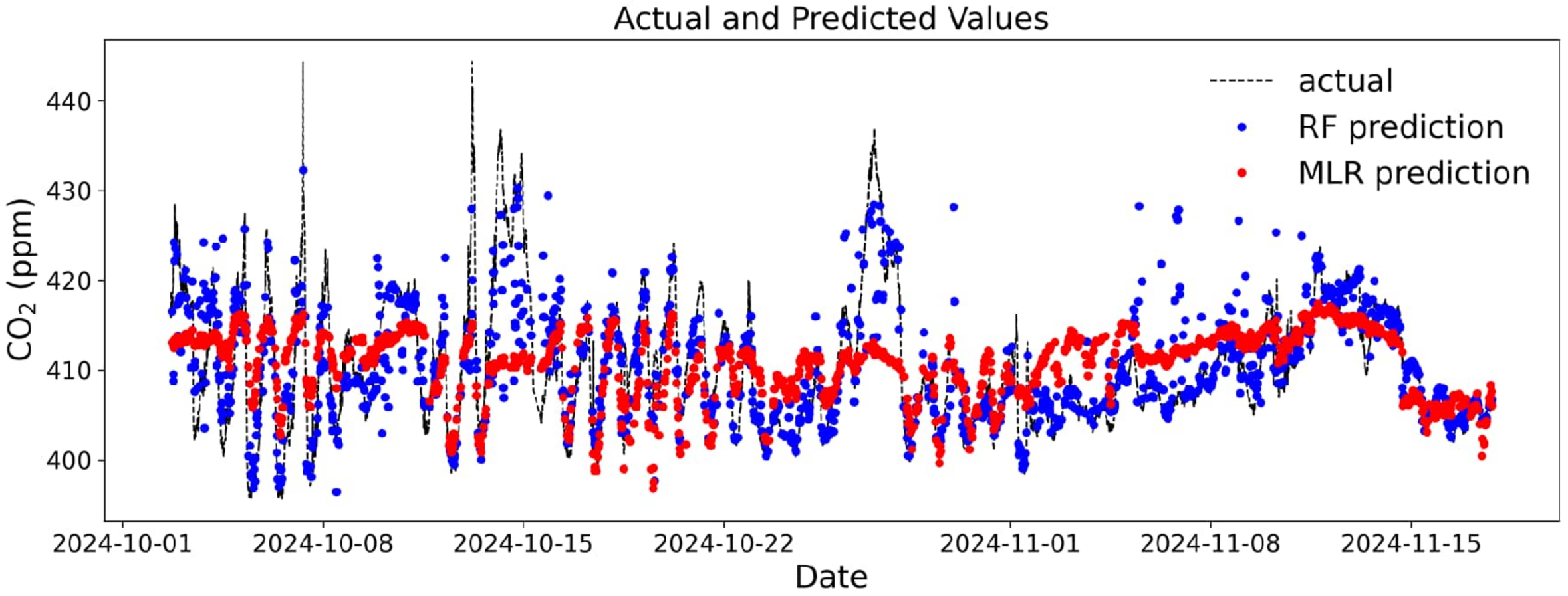


Comparing low-cost and high accuracy and cost sensor systems



Corrected low cost sensor signals compared to calibrated LGR laser system

Random forest classifier get the dynamic better

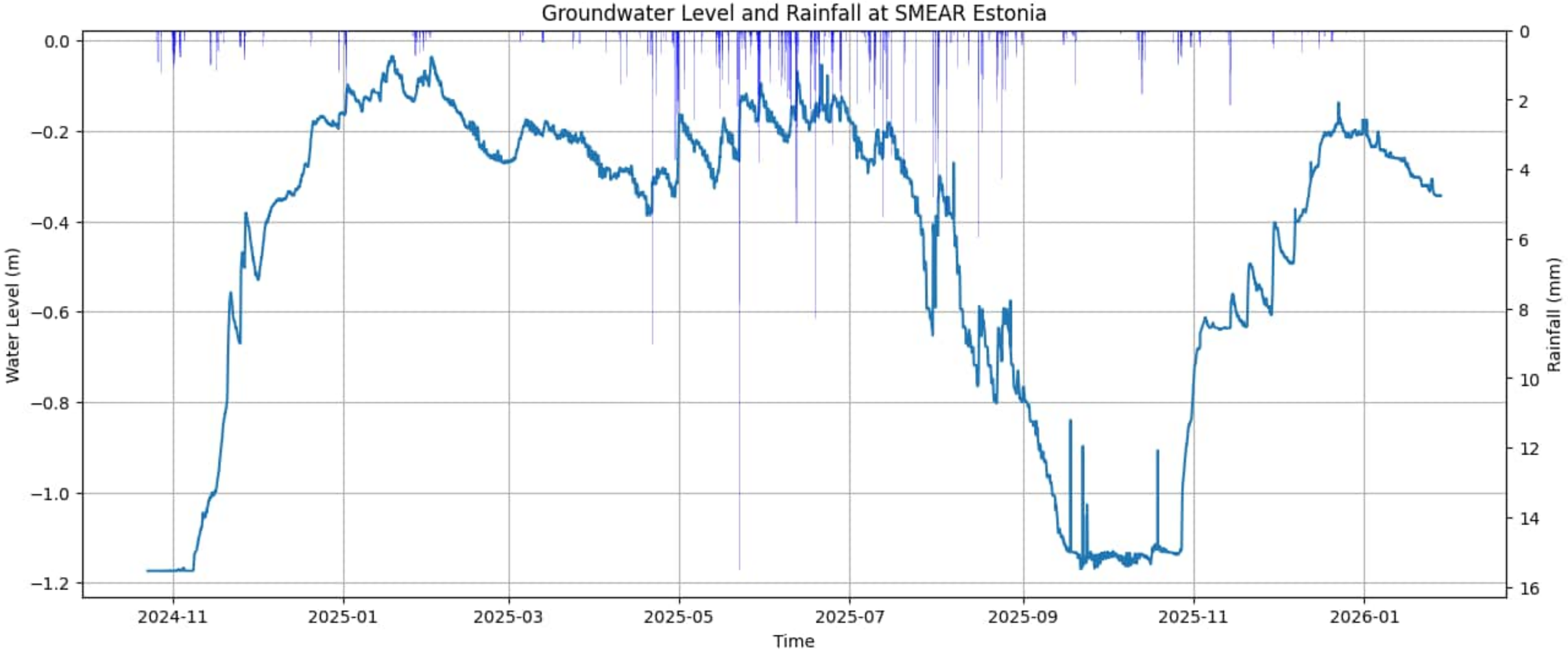


Groundwater table sensor

Pressure measurement to assess water level above the first gley level at ~1.2 m depth

Groundwater time series since October 2024

This enables us at SMEAR Estonia to get high quality hydrology data and models of the catchment.



The new AI/ML cluster

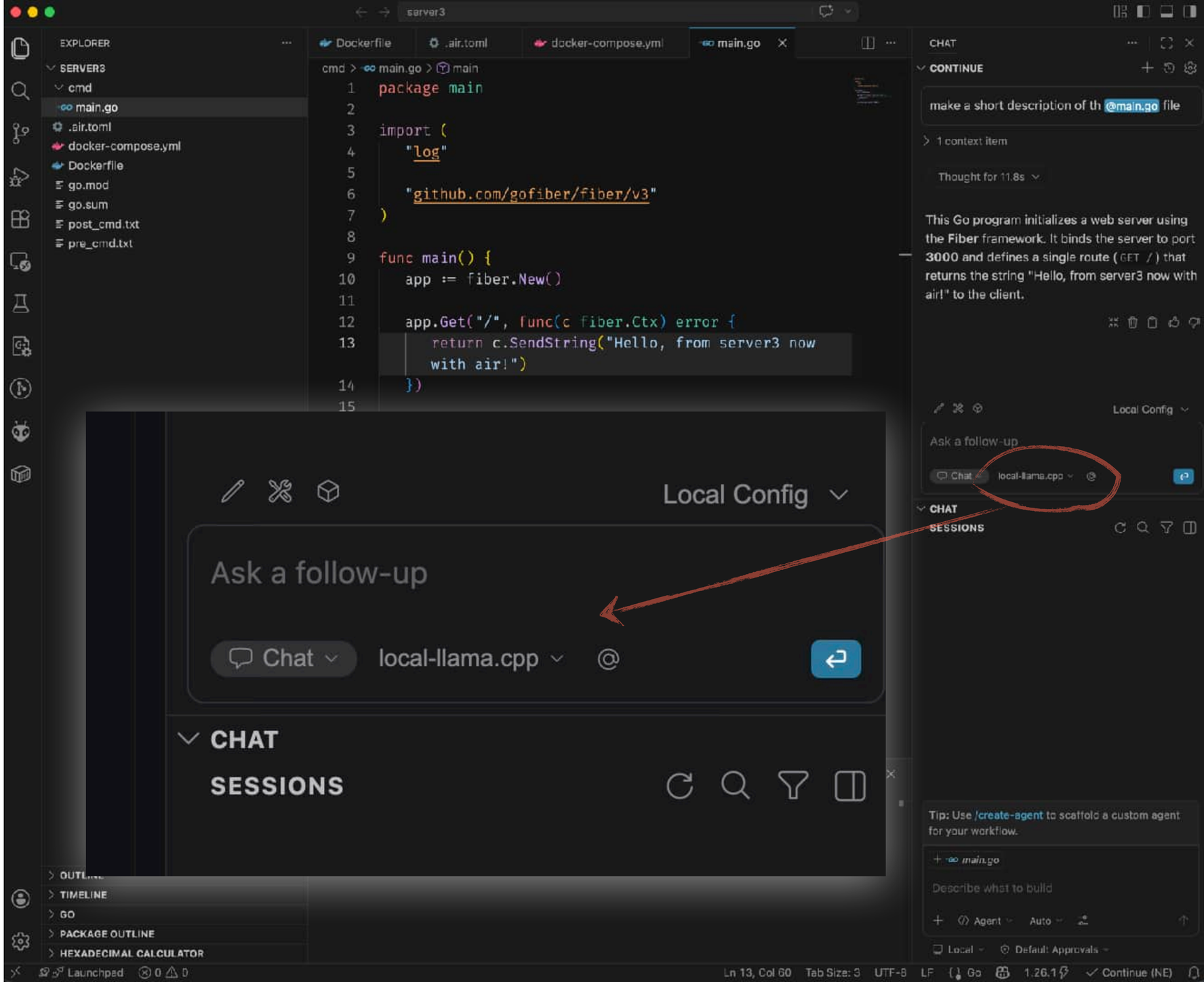
run and train local LLMs, ANNs from SMEAR data

- 3x 128 (384) GB unified RAM
- currently 4/1/1 (6) TB local SSD
- CPU: AMD RYZEN AI MAX+ PRO 395 (32) @ 5.19 GHz
- GPU: AMD Radeon 8060S Graphics
- System: CachyOS desktop (master) and Fedora 43 server (slaves)
- access to the SMEAR server and NAS in Tartu

Self hosted LLMs

Help in code development and education.

- We currently run on the AMD platform containerized LLM models
- gpt-oss-120b-Q4-K-M for multi purpose and agent mode
- GLM-4.7-Flash-Q8 for code generation works also in agent mode
- We use llama-cpp as server for both models
- They can be linked with code editors like VSCode or zed
- We can use them via web chat
- We can fine-tune the local AI models with our data



Paradigm shifts

- With LLM / AI support it is not anymore critical to know one tool / programming language very well. Standard tasks can be delegated to the AI.
- The need to understand the concepts is growing, then AI can be of help.
- Self-hosting model with hundreds of billions of parameters is fact, eg, each of the HP servers can host models up to 124 GB memory usage - and that at a price range of ~3 - 4 kEuro.
- Self-hosted fine-tuned models can be very trustworthy in information retrieval compared to the multipurpose model of the big players. Important in education!
- Low cost sensor development and AI let us increase the number of possible sensors per area and get well enough calibration for scientific and educational tasks.

Thanks

emu.ee

maaylikool

maaylikool

Eesti Maaülikool