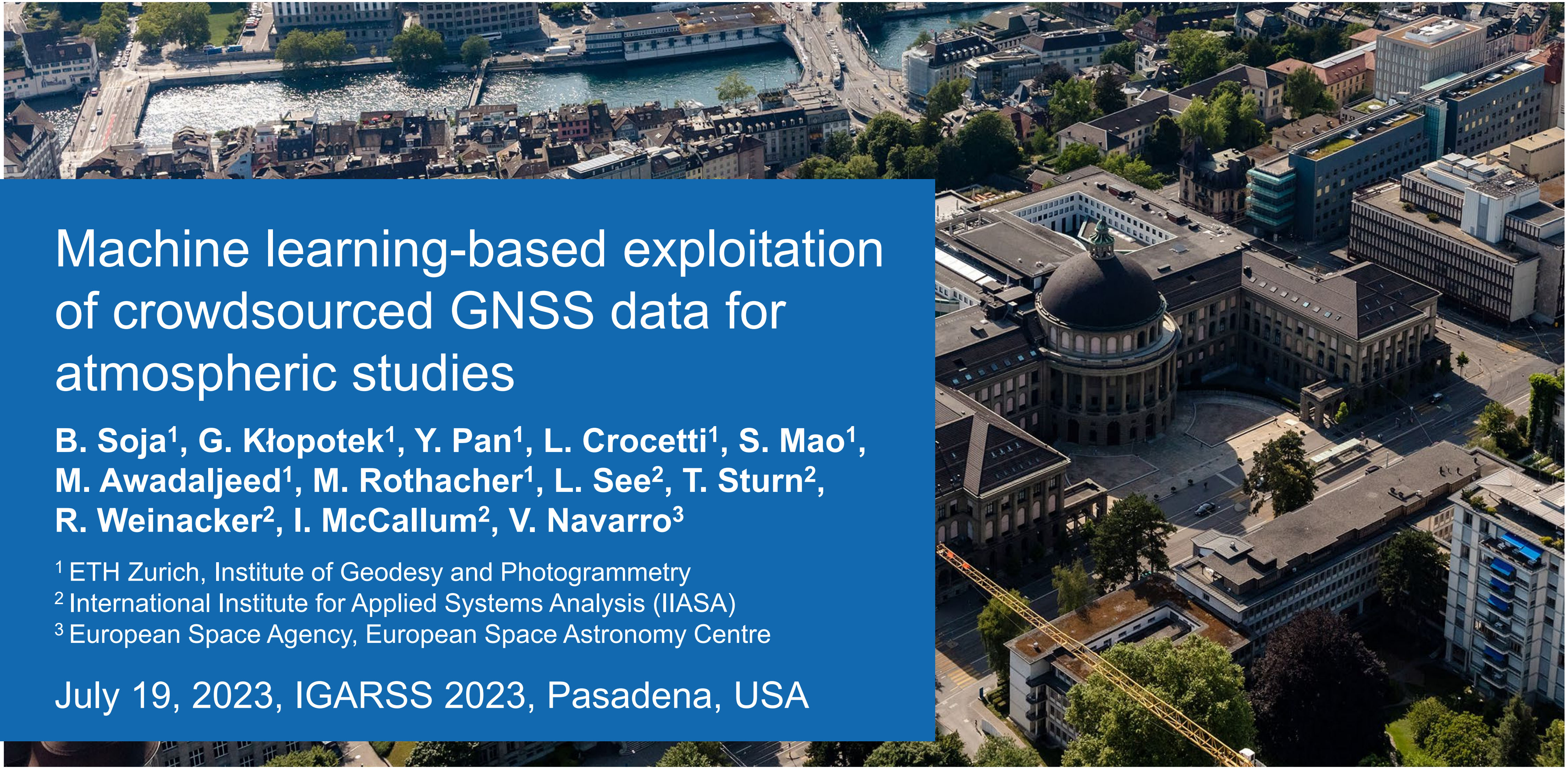
Machine learning-based exploitation of crowdsourced GNSS data for atmospheric studies

Author(s):
Soja, Benedikt; Klopotek, Grzegorz; Pan, Yuanxin; Crocetti, Laura; Mao, Shuyin; Awadaljeed, Mudathir; Rothacher, Markus; See, Linda; Sturn, Tobias; Weinacker, Rudi; McCallum, Ian; Navarro, Vicente

Publication date:
2023

Permanent link:
https://doi.org/10.3929/ethz-b-000624904



Machine learning-based exploitation of crowdsourced GNSS data for atmospheric studies

B. Soja[1], G. Kłopotek[1], Y. Pan[1], L. Crocetti[1], S. Mao[1], M. Awadaljeed[1], M. Rothacher[1], L. See[2], T. Sturn[2], R. Weinacker[2], I. McCallum[2], V. Navarro[3]

[1] ETH Zurich, Institute of Geodesy and Photogrammetry
[2] International Institute for Applied Systems Analysis (IIASA)
[3] European Space Agency, European Space Astronomy Centre

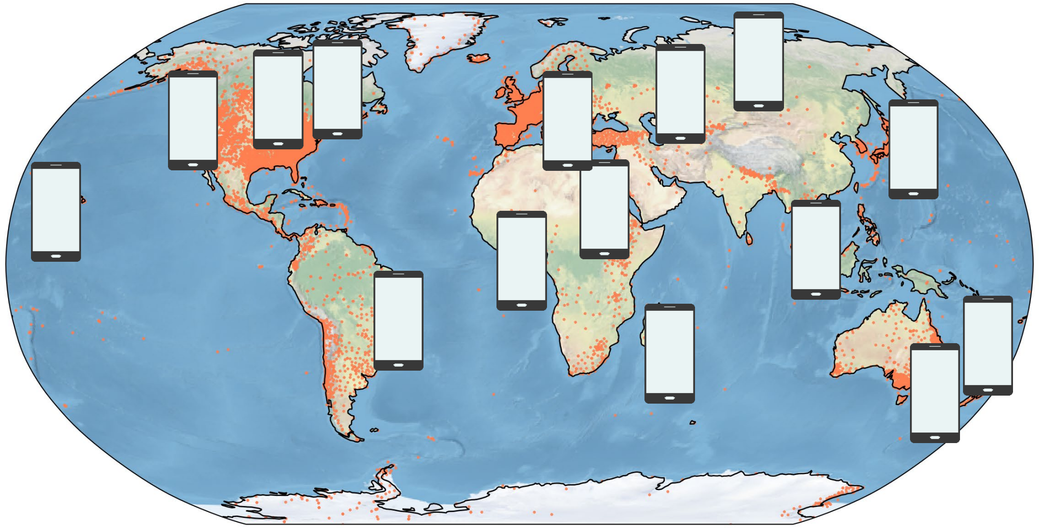
July 19, 2023, IGARSS 2023, Pasadena, USA

Motivation

- GNSS is capable of sensing atmospheric properties
  - Water vapor (troposphere)
  - Electron content (ionosphere)


- Models of the atmosphere beneficial for
  - Improving understanding of atmospheric processes
  - (Space) weather forecasts
  - Correcting positioning errors of low-cost devices

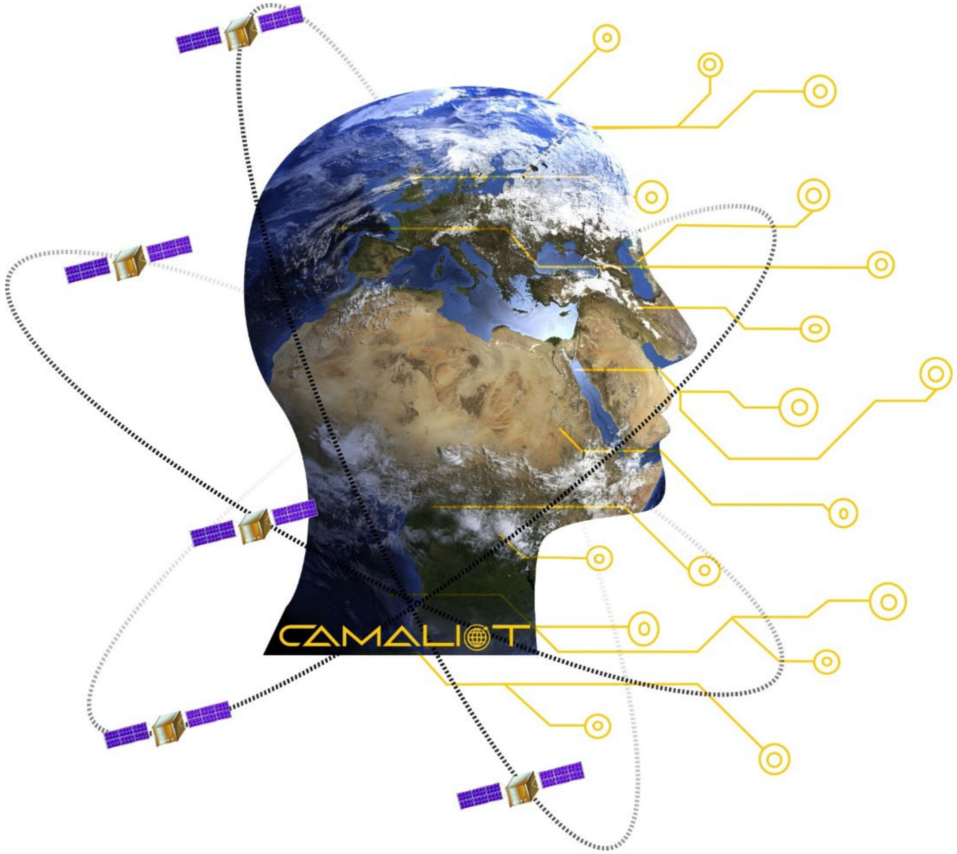
- GNSS big data: potential for machine learning (ML)
  - 20k geodetic stations
  - Billions of GNSS-capable devices



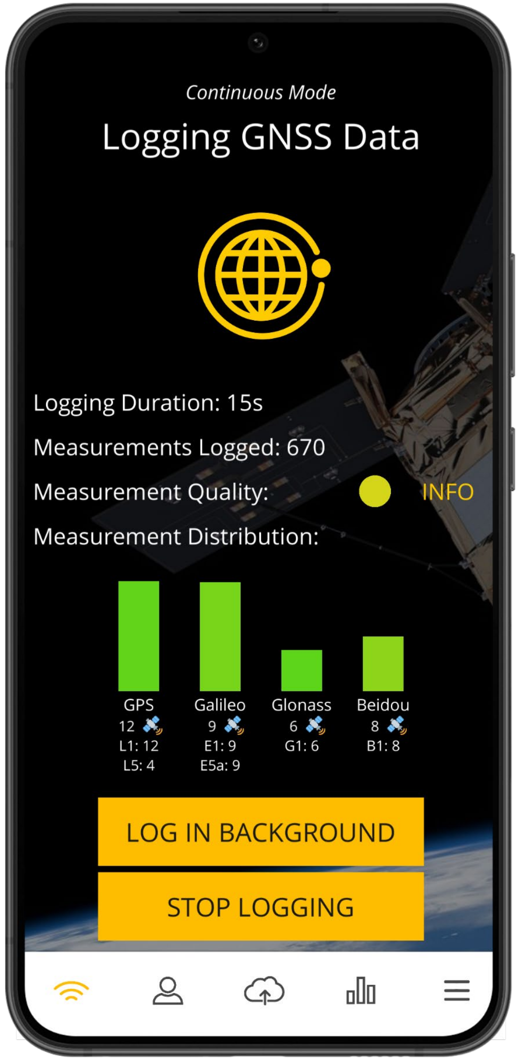
CAMALIOT

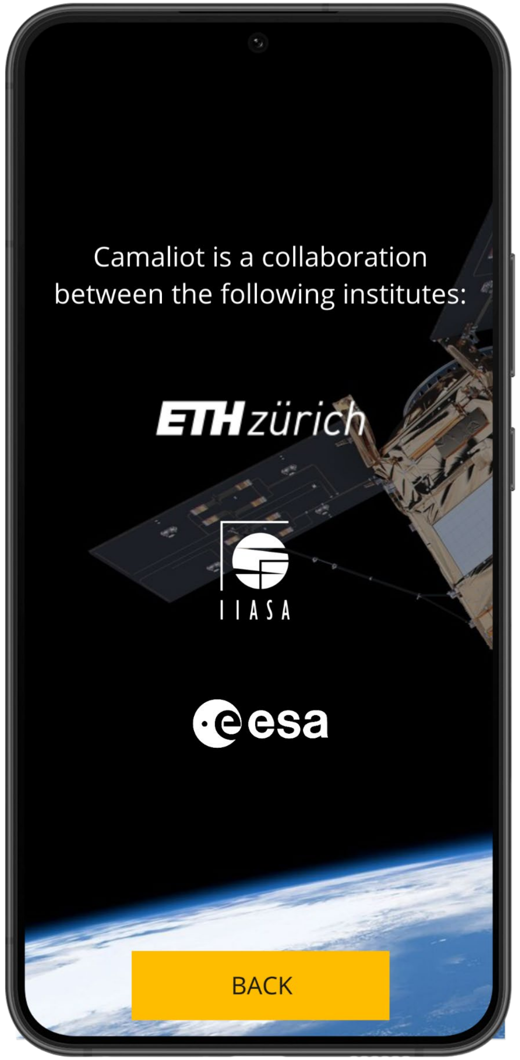
Application of machine learning technology for GNSS IoT data fusion

https://camaliot.org



CAMALIOT mobile app



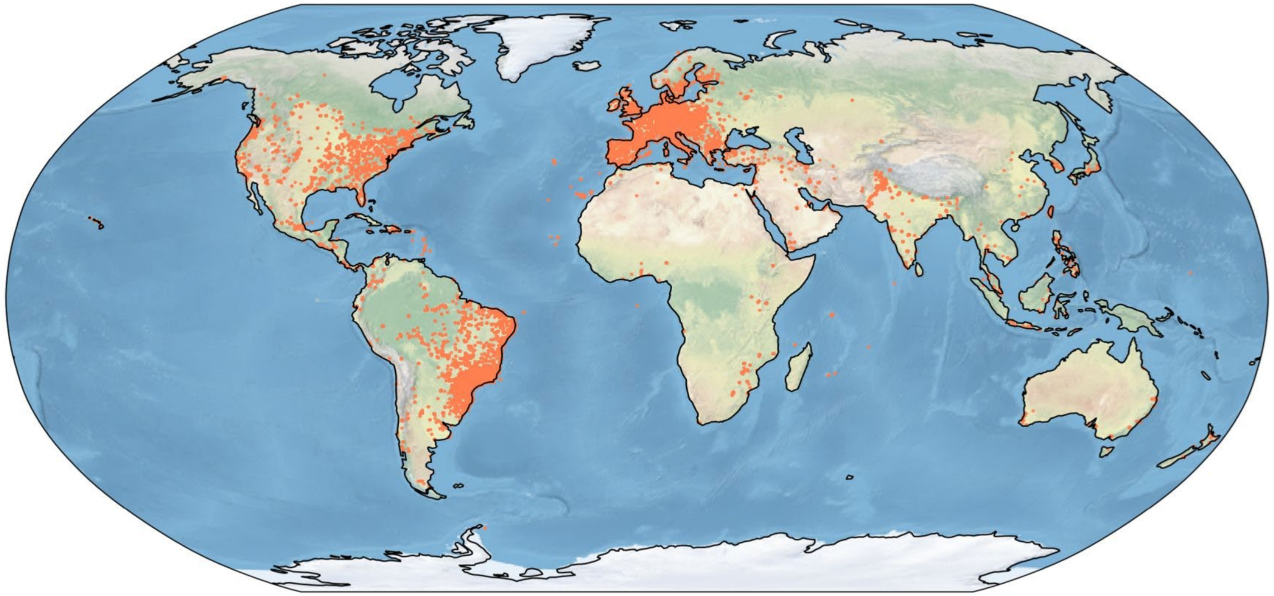


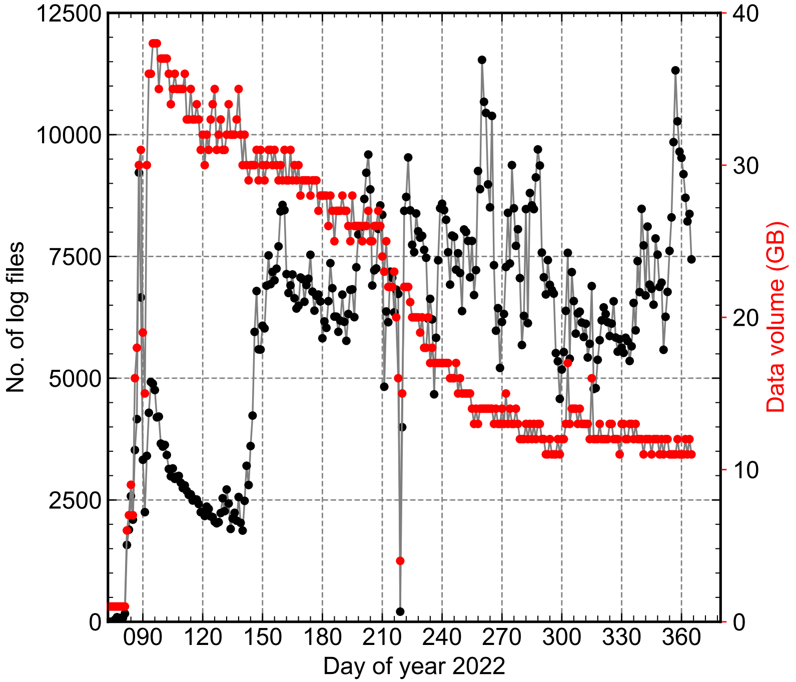
CAMALIOT crowdsourcing campaigns

- First campaign
  - March 17th – July 31st, 2022
  - 12k users
  - 116 billion GNSS measurements
- Autumn campaign
  - August 1st – November 30th, 2022
  - 1.2k users
  - 40 billion GNSS measurements

See et al., 2023, Int. J. Digit. Earth

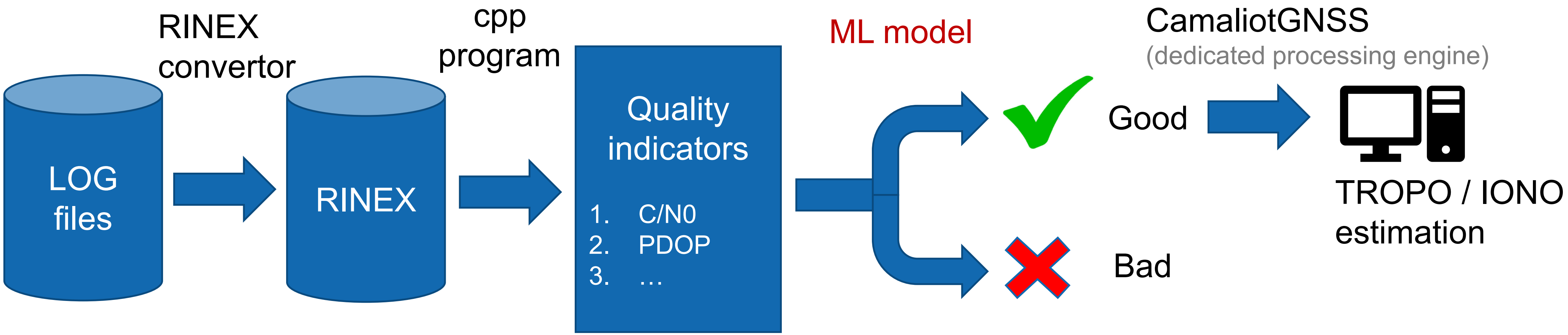
Data collection ongoing!





Smartphone GNSS data quality characterization

Machine learning-based data selection

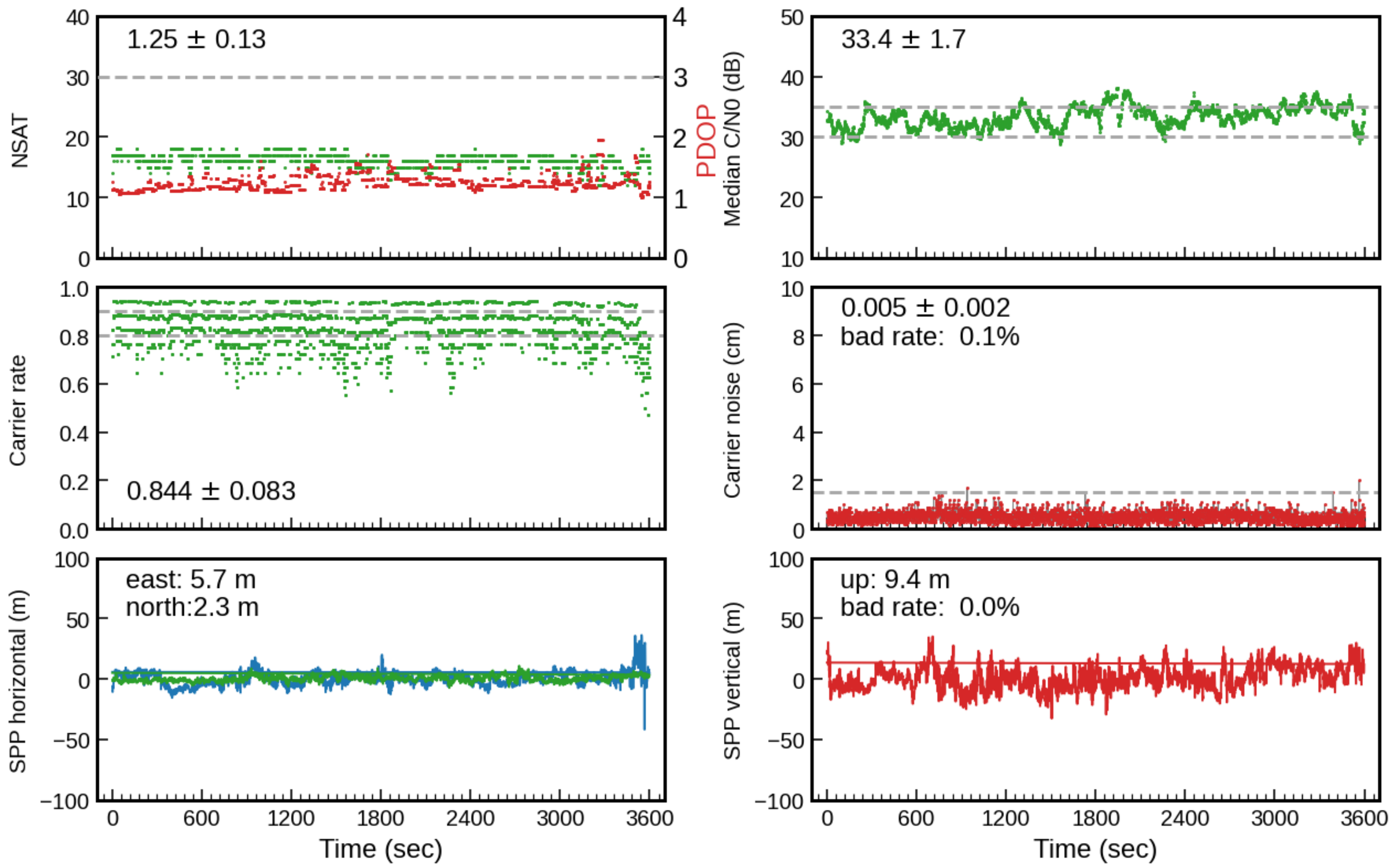


Pan et al., D4G 2022
& paper in preparation

Data labeling

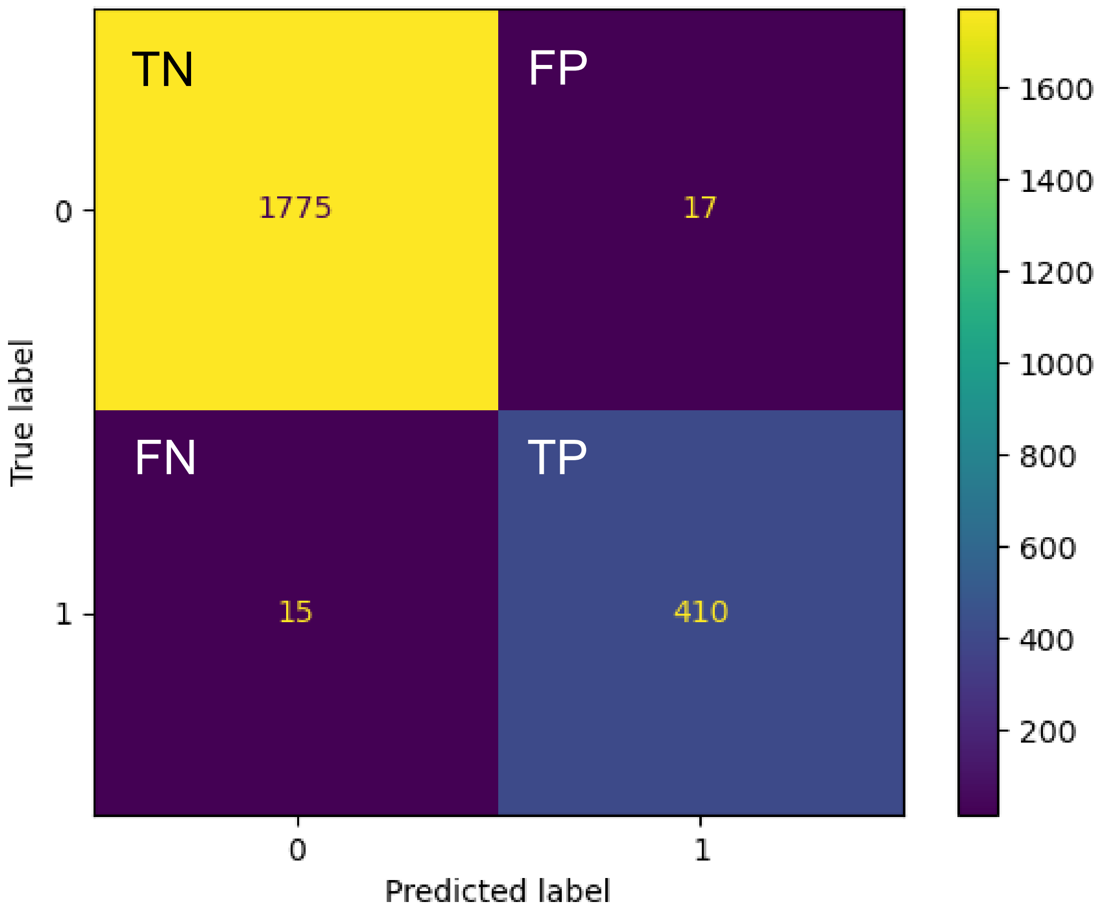
- Generate quality indicators
  - Number of epochs and satellites
  - PDOP
  - Carrier-to-noise ratio
  - SPP precision and solution availability
  - Time-differenced carrier phase precision and solution availability

- Labeling
  - Automatically reject very bad data
  - Final labeling carried out by human



Machine learning model

- Algorithm: random forest
- Training dataset
  - DOY 100-130: 31 days of hourly data
  - Good data: 1700
  - Bad data: 3400
- Training method
  - Grid search for hyperparameter tuning
  - 4-fold cross validation
- Test dataset
  - Good data: 425
  - Bad data: 1792



0.96 recall (TP/(TP+FN))

Smartphone data for atmospheric modeling

- Processing of “good” crowdsourced smartphone data still challenging → done manually
  - Large observation noise
  - Frequent occurrence of cycle slips
- Currently only relative atmospheric information
  - Differential tropospheric delays w.r.t. geodetic stations
  - No receiver differential code bias estimation yet for ionospheric parameters

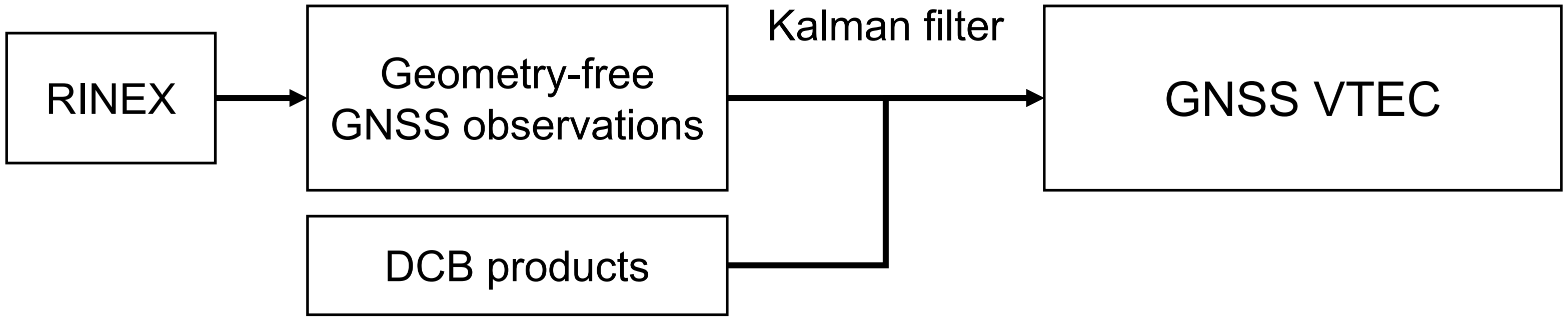
Not yet possible to integrate smartphone observations into ML-based models of the atmosphere

→ ML-based atmospheric models used for validation

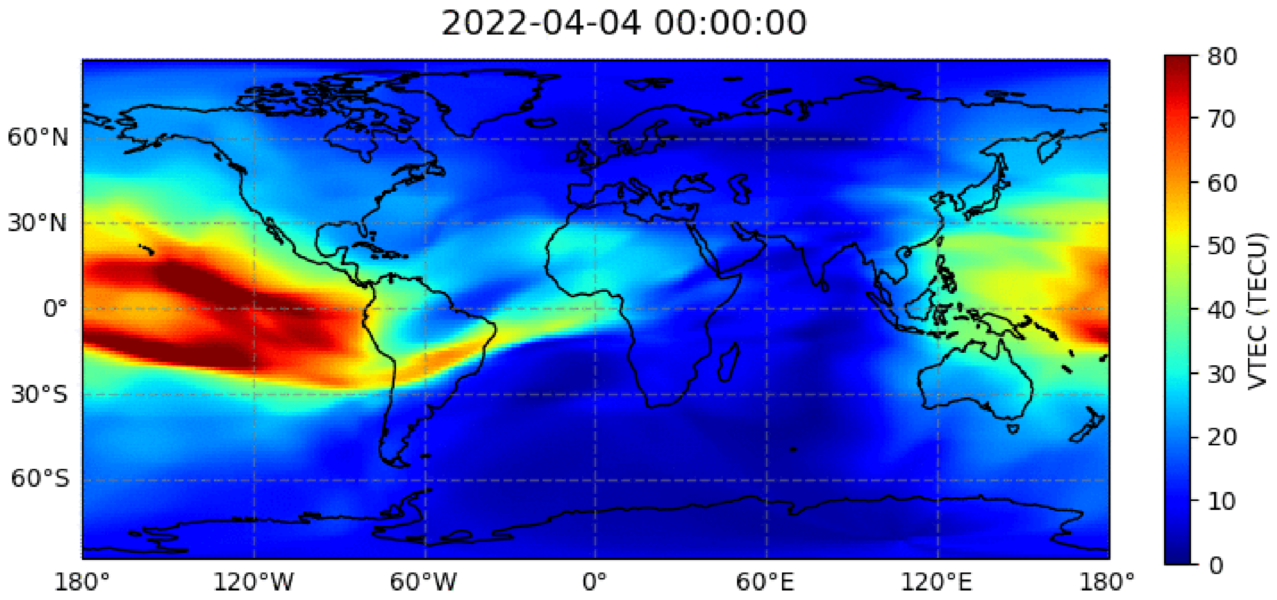
ML-based modeling of the ionosphere

Mao et al., EGU23
& paper in preparation

- Spatial modeling of vertical total electron content (VTEC)
- VTEC determination:



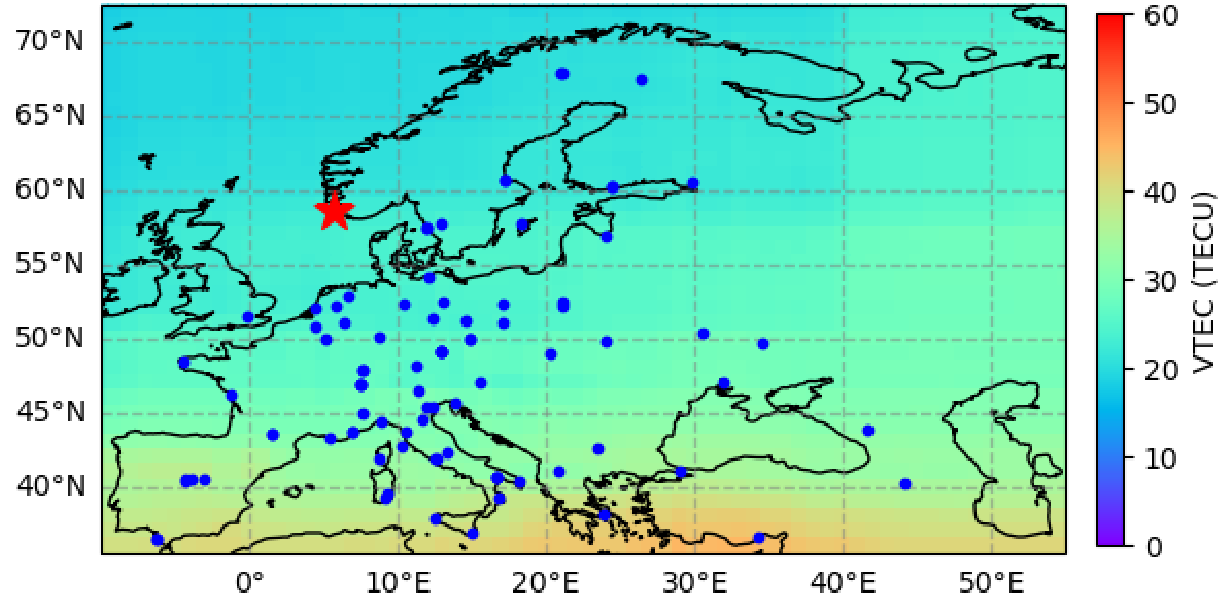
- ML model:
  - Input: latitude, longitude, time
  - Target: VTEC
  - Algorithm: random forest



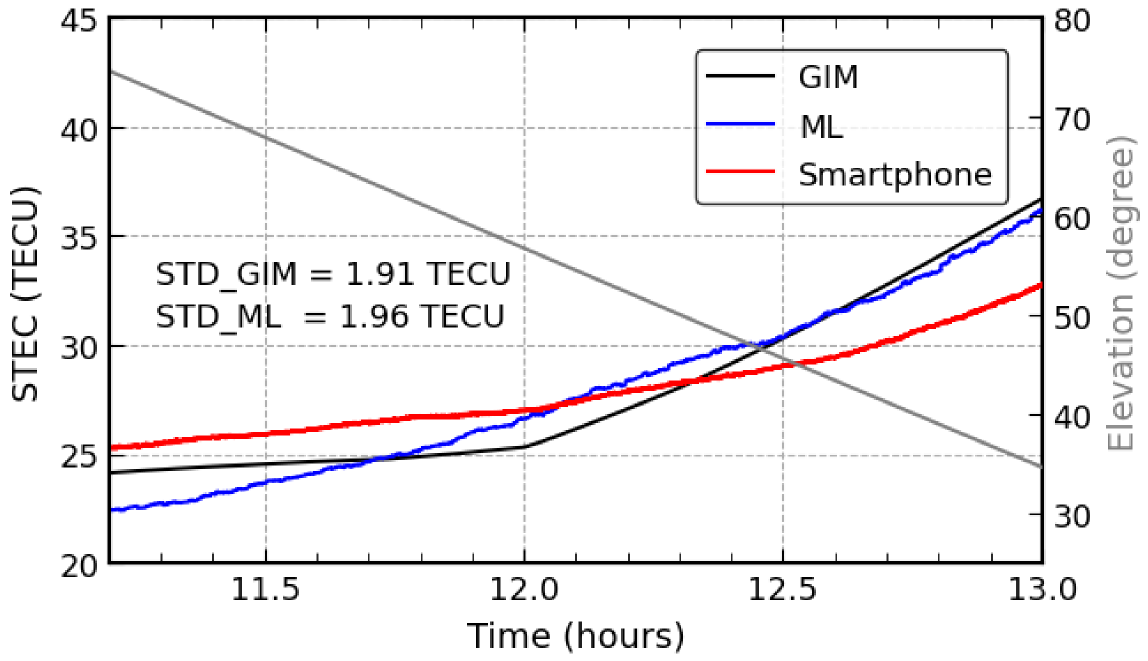
Example smartphone results for ionosphere

- Comparison of slant total electron content (STEC) for Galileo satellite E02 on April 4, 2022, from:
  - Smartphone (rDCB set manually)
  - IGS global ionospheric map
  - ML model

Location of smartphone, VTEC from ML model



STEC time series comparison

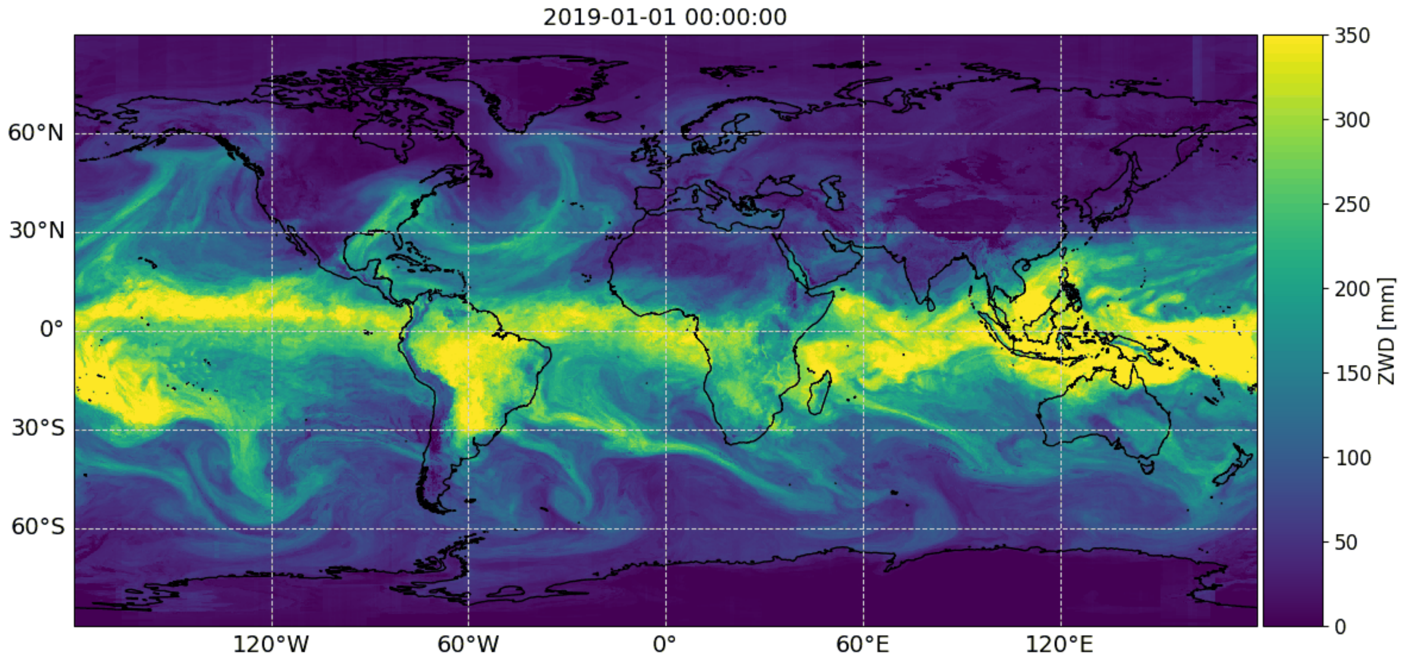


ML-based modeling of the troposphere

Crocetti et al., submitted to J Geod

- Spatial modeling of zenith wet delay (ZWD)
- Incorporation of meteorological parameters from numerical weather model

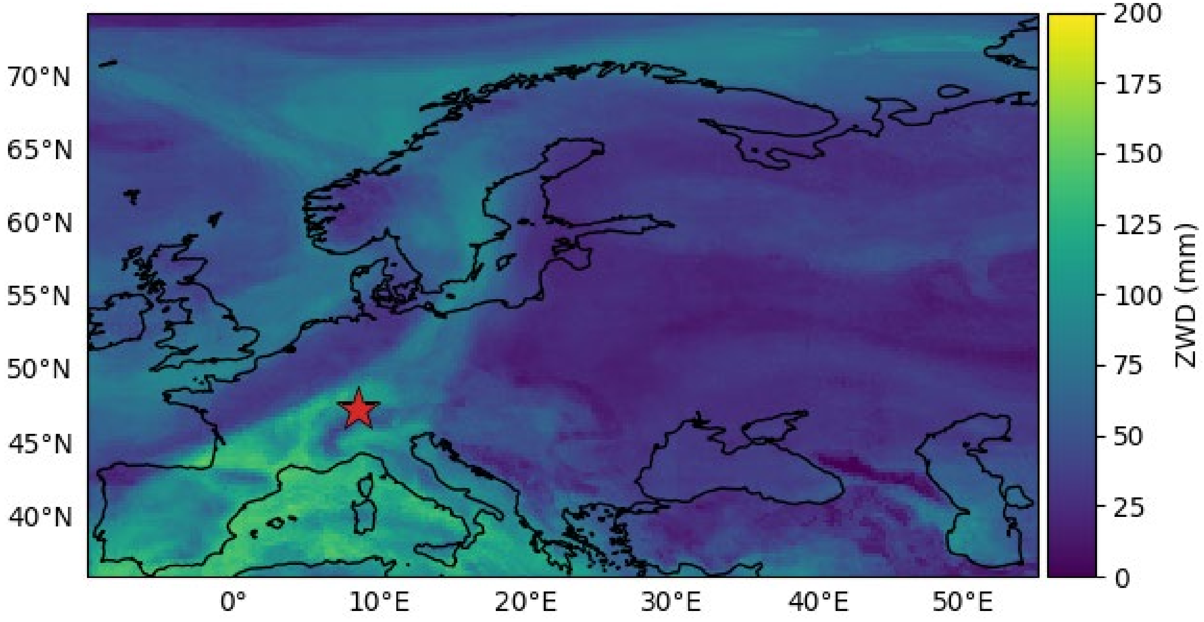
- ML model:
  - Input: latitude, longitude, height, time, specific humidity at 6 pressure levels (ECMWF)
  - Target: ZWD (NGL)
  - Algorithm: XGBoost



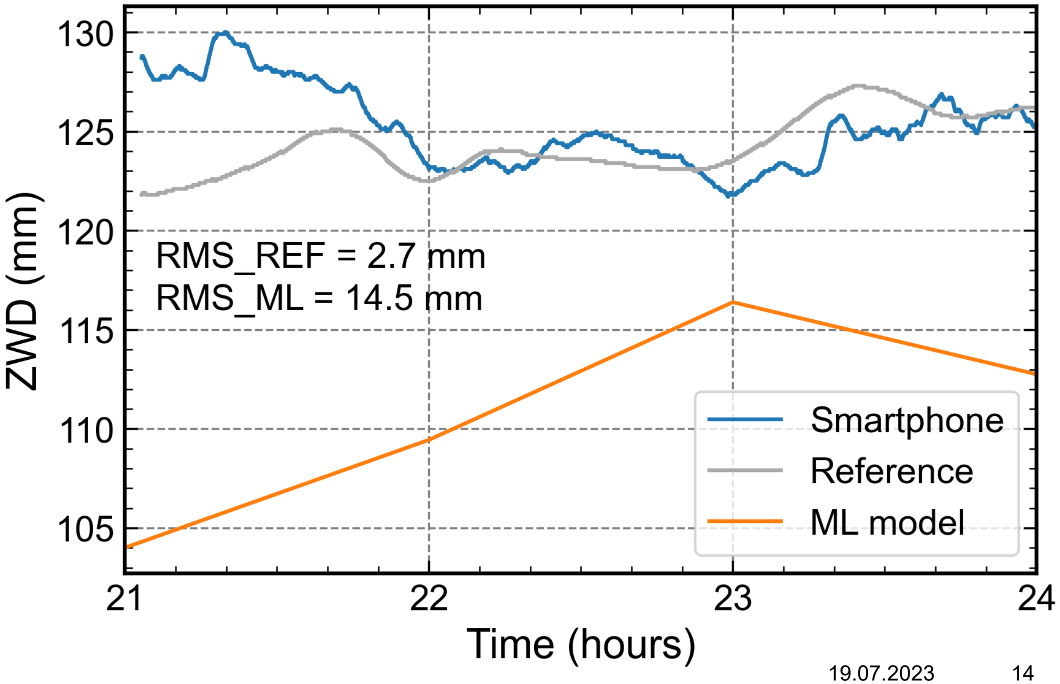
Example smartphone results for troposphere

- Comparison of ZWD on March 17, 2022, from:
  - Smartphone (differential ZWD determined and added to that of geodetic station ETH2)
  - Reference GNSS station located nearby (height difference considered)
  - ML model

Location of smartphone, ZWD from ML model



ZWD time series comparison



Summary

- Crowdsourced **smartphone GNSS** observations for atmospheric monitoring
- **ML applied** for
  - Quality characterization
  - Ionospheric modeling
  - Tropospheric modeling
- Smartphone data processing challenging, but **examples of high quality**

Kłopotek et al., submitted to Adv. Space Res.

IGARSS proceedings paper

https://camaliot.org

ETH zürich

SPACE GEODESY

Prof. Dr. Benedikt Soja

soja@ethz.ch

@SpaceGeodesyETH

@b_soja

ETH Zurich
Chair of Space Geodesy
HPV G 54
Robert-Gnehm-Weg 15
8093 Zurich, Switzerland

www.space.igp.ethz.ch

Thanks for your attention!