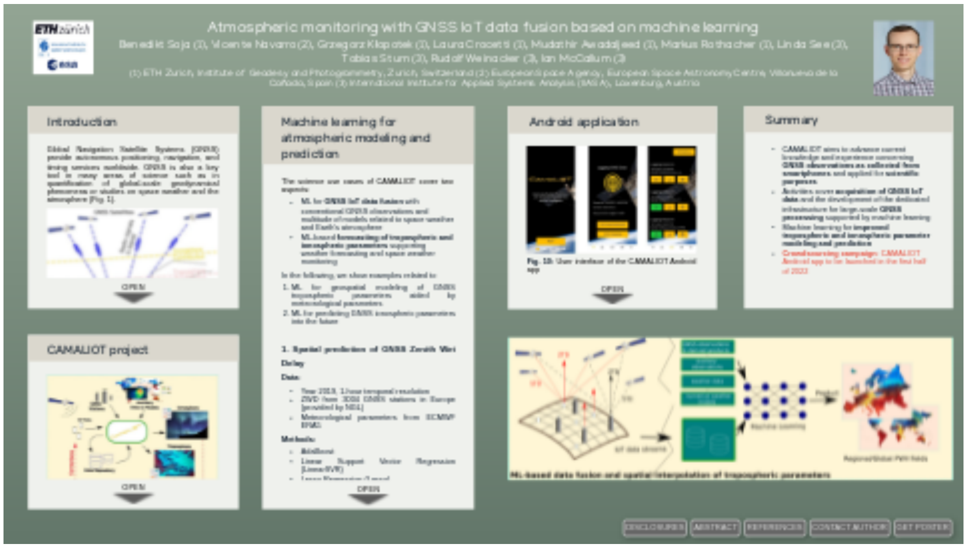
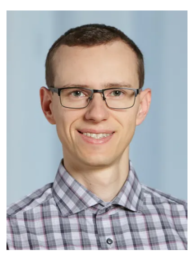
Atmospheric monitoring with GNSS IoT data fusion based on machine learning



Benedikt Soja (1), Vicente Navarro (2), Grzegorz Kłopotek (1), Laura Crocetti (1), Mudathir Awadaljeed (1), Markus Rothacher (1), Linda See (3), Tobias Sturn (3), Rudolf Weinacker (3), Ian McCallum (3)

(1) ETH Zurich, Institute of Geodesy and Photogrammetry, Zurich, Switzerland (2) European Space Agency, European Space Astronomy Centre, Villanueva de la Cañada, Spain (3) International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria



PRESENTED AT:

INTRODUCTION

Global Navigation Satellite Systems (GNSS) provide autonomous positioning, navigation, and timing services worldwide. GNSS is also a key tool in many areas of science such as in quantification of global-scale geodynamical phenomena or studies on space weather and the atmosphere (Fig. 1).

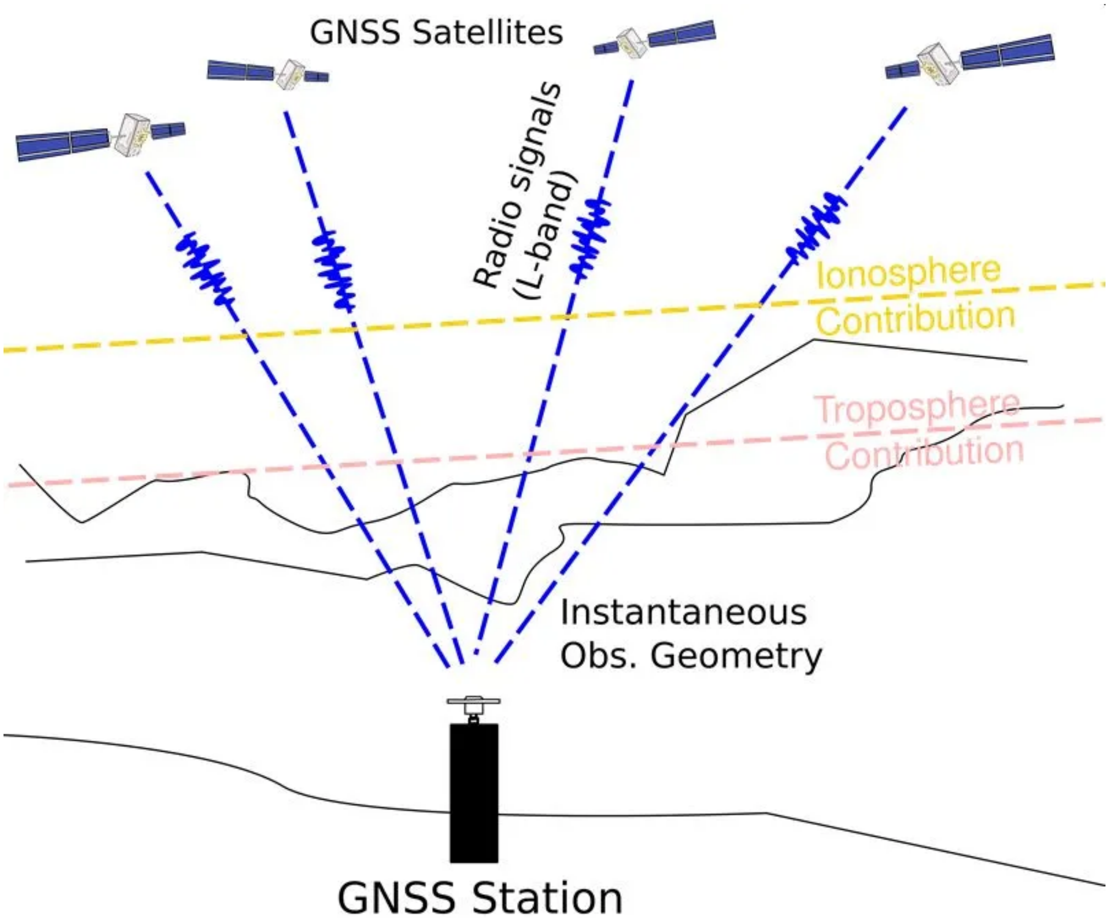


Fig. 1: GNSS measurements concerning the ionosphere and troposphere

Modern smartphones running on Android version 7.0 and higher enable the collection of raw GNSS measurements. The prospective utilization of a very large number of mobile phones with GNSS-capable receivers as a dynamic crowdsourcing receiver network would allow to significantly increase the spatio-temporal resolution available for the analysis of geodynamic and atmospheric phenomena.

The amount of GNSS data is growing continuously. Currently, data from around 20k geodetic-quality GNSS stations is publicly available (Fig. 2). Additionally, billions of Internet-of-things (IoT) devices with GNSS capabilities exist, highlighting the potential for crowdsourcing campaigns to collect such data. The large amount of data motivates the application of new concepts for data analysis, in particular machine learning (ML) approaches.

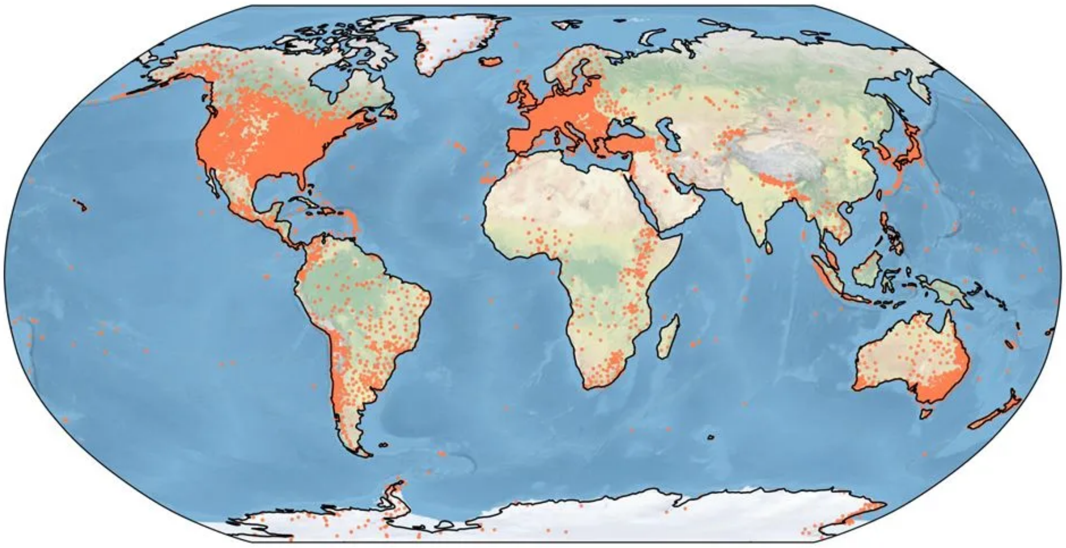


Fig. 2: Distribution of high-quality GNSS stations used for Earth observation

CAMALIOT PROJECT

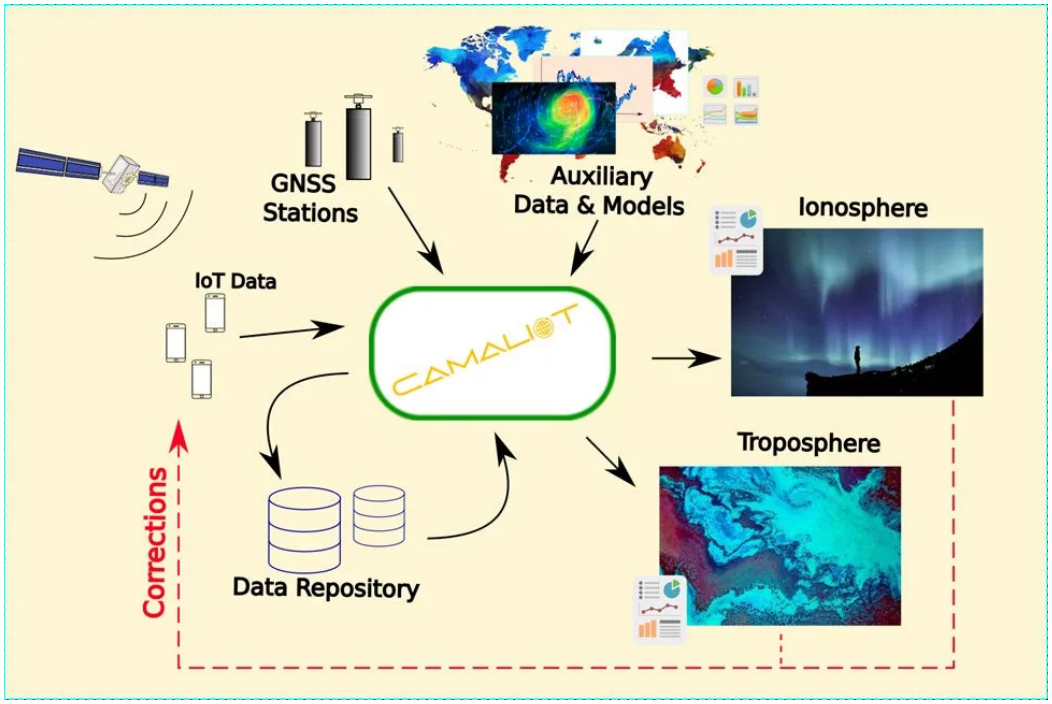


Fig. 3: High-level overview of CAMALIOT

Application of **Ma**chine **L**earning Technology for GNSS **IoT** data fusion (CAMALIOT) is an ongoing project funded by the European Space Agency (ESA) as part of the NAVISP Element 1 program. This activity is carried out by a consortium consisting of ETH Zurich and the International Institute for Applied Systems Analysis (IIASA). The main goals of CAMALIOT include:

- Development of an **Android application** and organization of a **crowdsourcing campaign** to acquire large amounts of GNSS observations from hundreds of smartphones (offline and real time)
- Establishment of a dedicated infrastructure for **GNSS processing and machine learning** at scale
- Using machine learning for the fusion of collected GNSS IoT data with geodetic GNSS data and external models to **improve models and predictions of the troposphere and ionosphere**

MACHINE LEARNING FOR ATMOSPHERIC MODELING AND PREDICTION

The science use cases of CAMALIOT cover two aspects:

- ML for **GNSS IoT data fusion** with conventional GNSS observations and multitude of models related to space weather and Earth's atmosphere
- ML-based **forecasting of tropospheric and ionospheric parameters** supporting weather forecasting and space weather monitoring

In the following, we show examples related to

1. ML for geospatial modeling of GNSS tropospheric parameters aided by meteorological parameters
2. ML for predicting GNSS ionospheric parameters into the future

1. Spatial prediction of GNSS Zenith Wet Delay

Data:

- Year 2019, 1-hour temporal resolution
- ZWD from 3004 GNSS stations in Europe (provided by NGL)
- Meteorological parameters from ECMWF ERA5

Methods:

- AdaBoost
- Linear Support Vector Regression (LinearSVR)
- Lasso Regression (Lasso)
- ElasticNet
- Stochastic Gradient Decent (SGD)
- Ridge Regression (Ridge)
- Multilayer Perceptron (MLP)
- Histogram-Based Gradient Boosting (HistGBoost)
- Extreme Gradient Boosting (XGBoost)
- Random Forest (RF)

Input features: latitude, longitude, time, meteorological data

Target feature: ZWD

80% of stations for training, 20% for testing

Results:

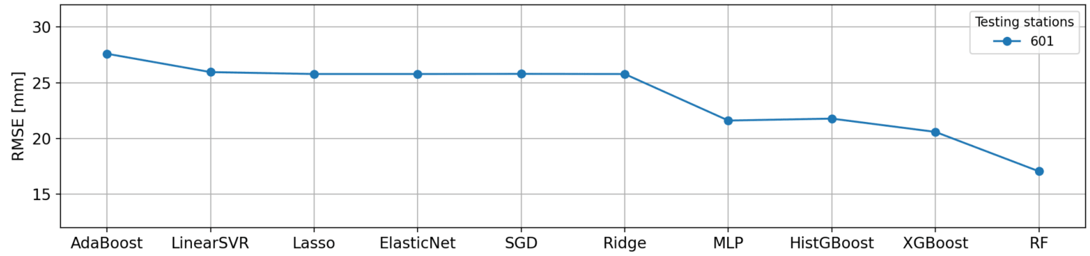


Fig. 4: Performance in terms of RMSE of different ML algorithms for spatial ZWD prediction assessed on the test set

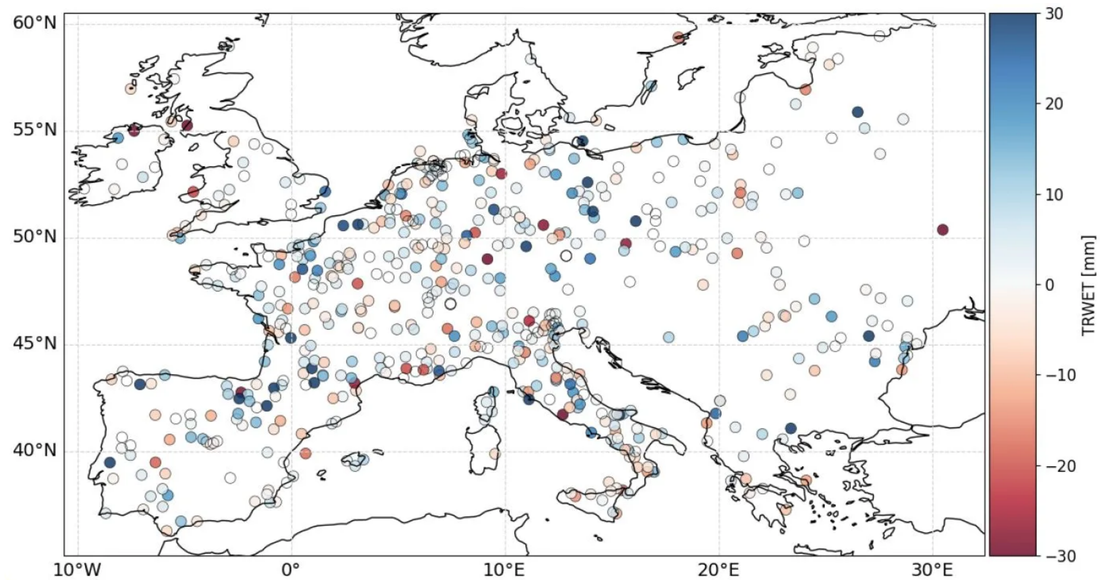


Fig. 5: Differences in ZWD between Random Forest predictions and actual observations for test stations on 2019/01/05 at 0 UTC.

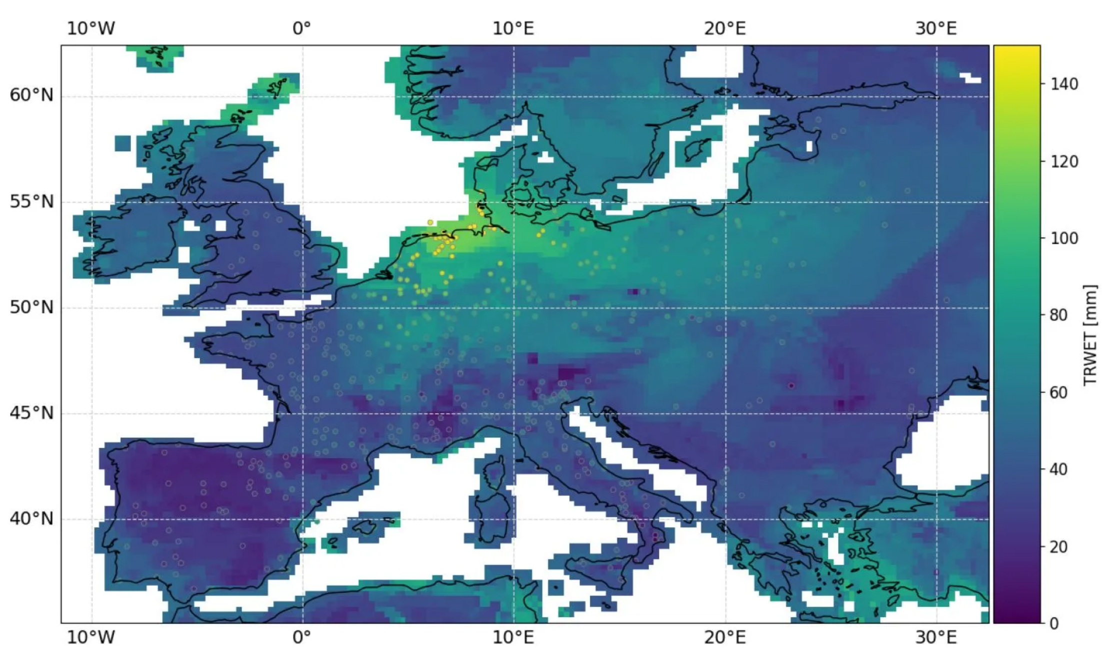


Fig. 6: Predicted ZWD based on the Random Forest model on 2019/01/05 at 0 UTC.

2. Temporal prediction of GNSS Vertical Total Electron Content

Data:

- Years 2002-2020, 2-hour temporal resolution
- VTEC from Global Ionospheric Map product of the International GNSS Service, 20 grid points over Europe
- Geomagnetic Kp index

Methods:

- Repeat values from previous day (baseline)
- Linear model
- Dense model (Artificial Neural Network)
- Convolutional Neural Network (CNN)
- Long Short-Term Memory (LSTM)
- Autoregressive LSTM (AR LSTM)

Input features: VTEC of previous 2 weeks, time (parameterized as sinusoids with daily and yearly periods)

Target feature: VTEC of next 24 hours

Time series split into training (2002-2015), validation (2015-2018) and testing (2019-2021).

Results:

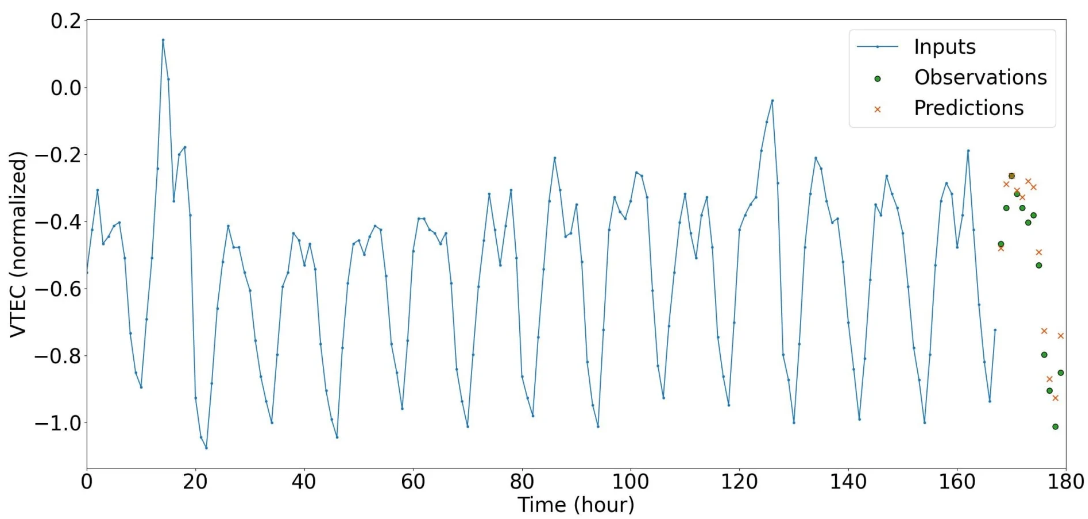


Fig. 7: Example of predicting VTEC 24 hours ahead at the location 50°N, 20°E with LSTM

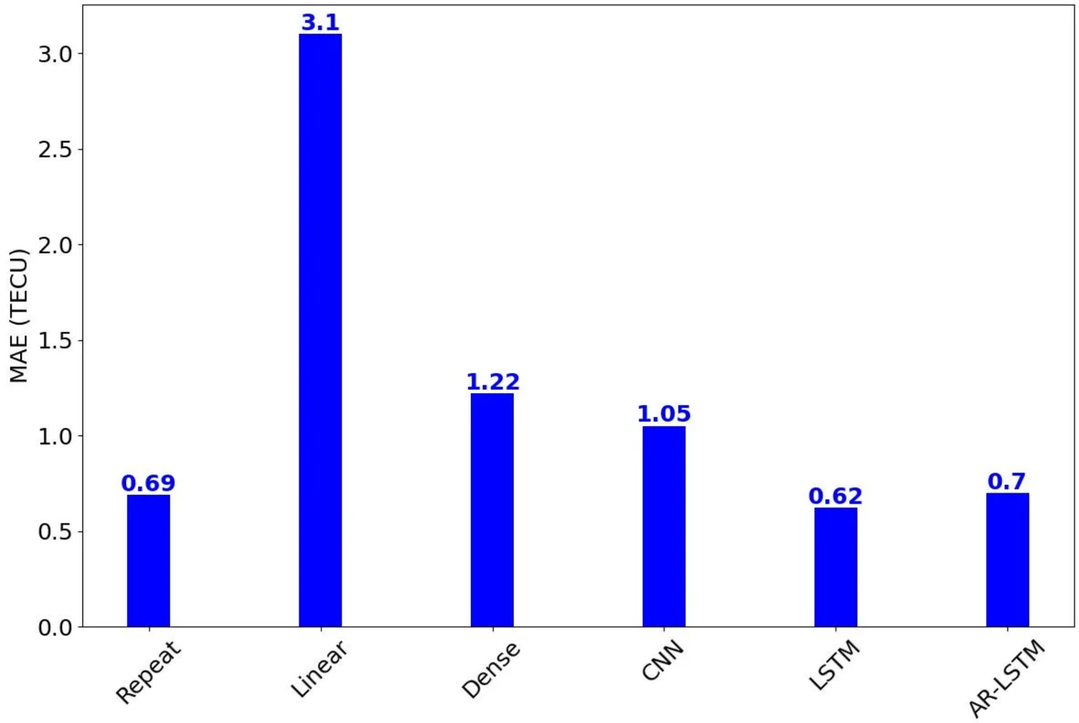


Fig. 8: Performance in terms of MAE for different ML algorithms for temporal VTEC prediction at location 50°N, 20°E assessed on the test set

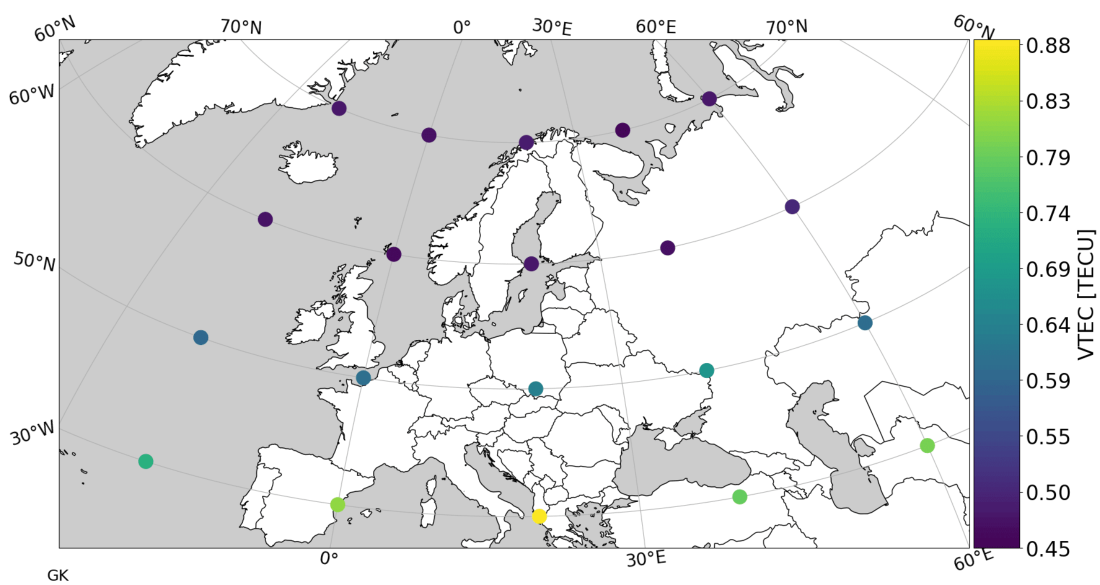


Fig. 9: Color-coded map representing the MAE of VTEC predictions on the test set for 20 grid points in Europe

ANDROID APPLICATION

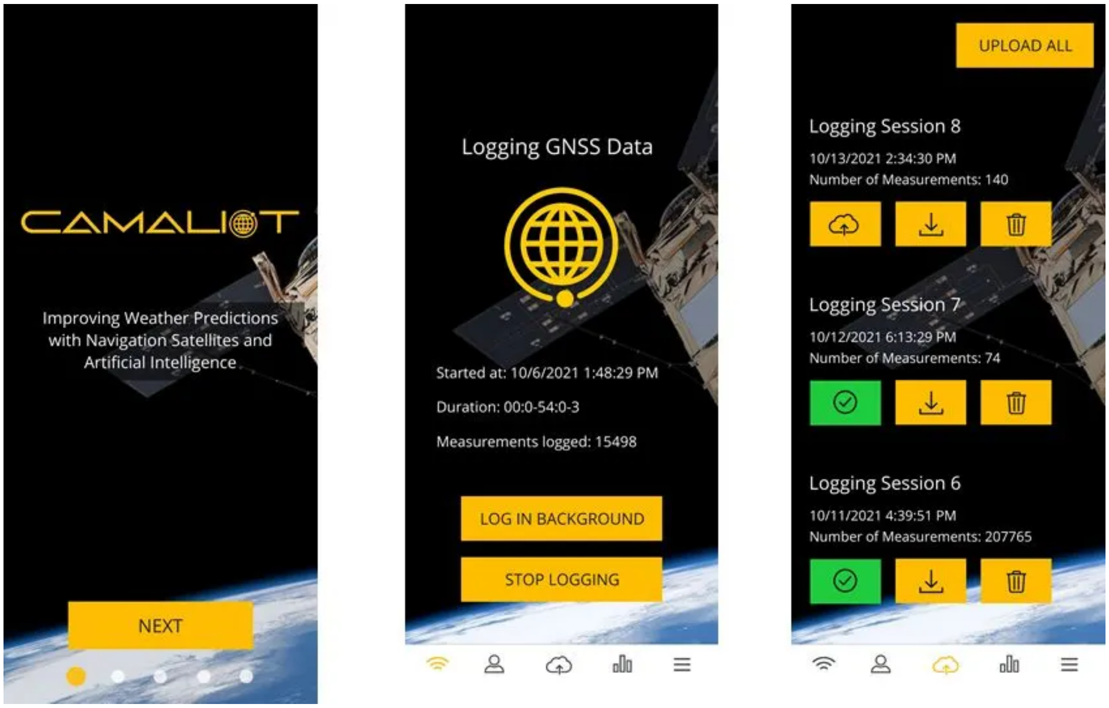


Fig. 10: User interface of the CAMALIOT Android app

The CAMALIOT Android app (Fig. 10) is capable of recording multi-frequency, multi-constellation raw GNSS data on recent smartphones. The user can download the raw GNSS data in RINEX format for his/her own purposes, while being encouraged to share the data by uploading it to the CAMALIOT servers. Elements of gamification (leaderboard) are designed to further increase the participation.

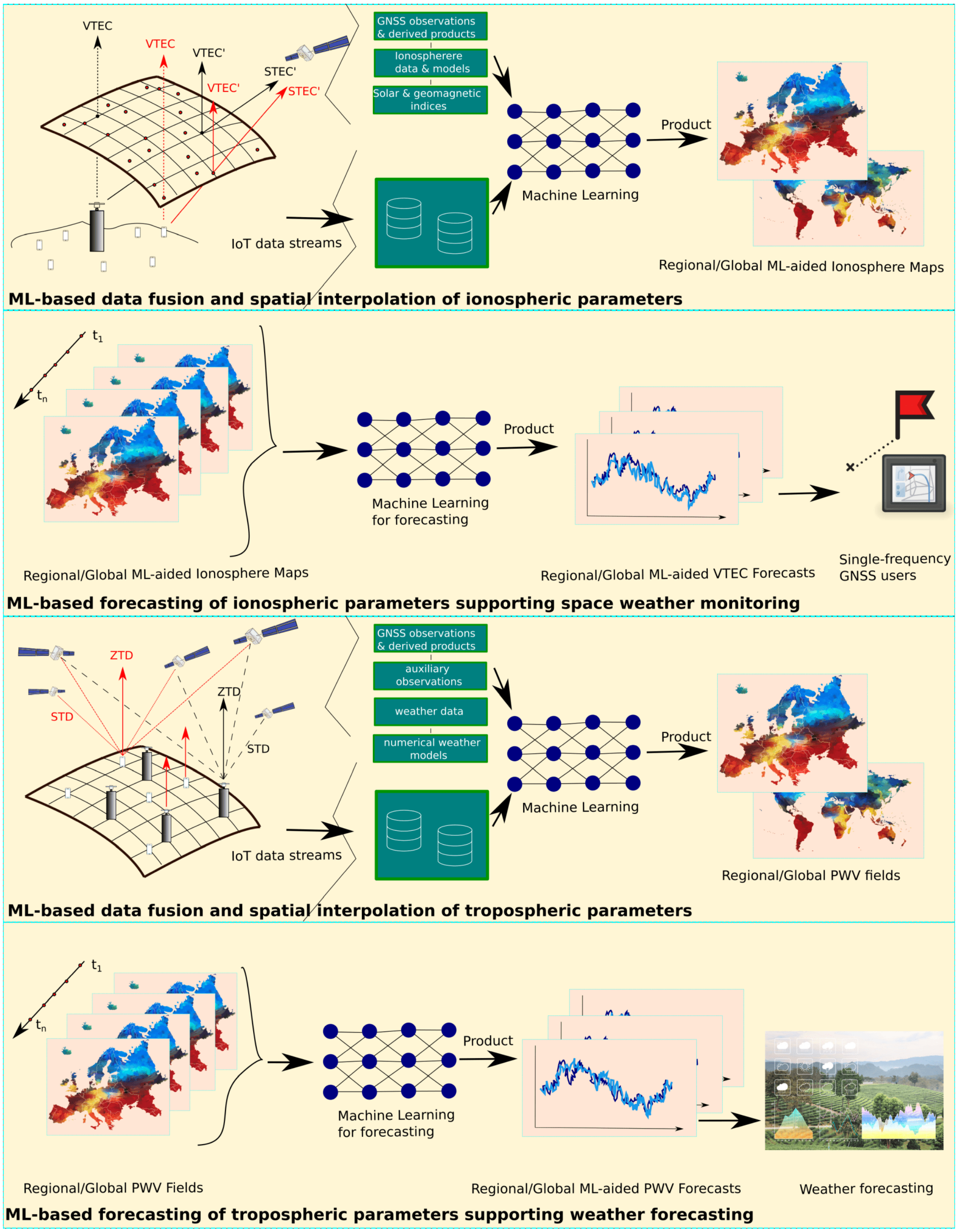
The app will be officially launched in the first half of 2022 as part of a crowdsourcing campaign to collect GNSS IoT data.

The video below shows a demonstration of logging GNSS data with the CAMALIOT app.

[VIDEO] https://res.cloudinary.com/amuze-interactive/video/upload/vc_auto/v1638891930/agu-fm2021/AE-AD-68-E7-D7-EC-B1-1C-E6-D8-EC-10-02-9A-F2-60/Video/Screen_Recording_20211207-164035_Camaliot_tr9lt6.mp4

SUMMARY

- CAMALIOT aims to advance current knowledge and experience concerning **GNSS observations as collected from smartphones** and applied for **scientific purposes**
- Activities cover **acquisition of GNSS IoT data** and the development of the dedicated infrastructure for large-scale **GNSS processing** supported by machine learning
- Machine learning for **improved tropospheric and ionospheric parameter modeling and prediction**
- **Crowdsourcing campaign**: CAMALIOT Android app to be launched in the first half of 2022



DISCLOSURES

This work is funded by the NAVISP Element 1 program of ESA.

ABSTRACT

Recent improvements in GNSS capabilities of various Internet-of-Things (IoT) devices, including smartphones, have allowed for better PNT performance, as well as for new potential applications in geosciences. In particular, multi-constellation and multi-frequency receivers found in recent generations of smartphones bring great potential for GNSS science exploitation. However, access to IoT data for scientific purposes is currently limited, facing different data processing challenges.

The project CAMALIOT: Application of Machine Learning Technology for GNSS IoT Data Fusion addresses these issues in order to increase the usability of GNSS IoT data for scientific purposes. It encompasses the whole pipeline from raw GNSS IoT data collection and development of methods for efficient and automatic processing, to the final suitability demonstration for determination and prediction of atmospheric parameters. In this way, the project will extend the capabilities of the GNSS Science Support Centre (GSSC) of ESA, which offers GNSS data and processing services for various domains.

In order to collect raw GNSS IoT data, an Android app has been developed that will be the basis of a crowdsourcing campaign. The data ingestion, processing, and analysis are designed to be highly automated, robust, and scalable. Machine learning is used for several tasks in the processing scheme, including anomaly detection as well as data fusion and prediction. The GNSS IoT data is combined with geodetic GNSS data and external models and datasets related to the atmosphere and space weather to realize global grids of tropospheric parameters (zenith wet delays, gradients, precipitable water vapor) and ionospheric vertical total electron content.

This work is funded by the NAVISP Element 1 program of ESA.

REFERENCES

Navarro, Vicente, Grieco, Raffaella, Soja, Benedikt, Nugnes, Marco, Klopotek, Grzegorz, Tagliaferro, Giulio, See, Linda, Falzarano, Roberto, Weinacker, Rudolf, VenturaTraveset, Javier, "Data Fusion and Machine Learning for Innovative GNSS Science Use Cases," *Proceedings of the 34th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2021)*, St. Louis, Missouri, September 2021, pp. 2656-2669. https://doi.org/10.33012/2021.18115 (https://doi.org/10.33012/2021.18115)