1. INTRODUCTION

Global Navigation Satellite System (GNSS) is mostly known for autonomous global positioning, navigation, and timing services. However, it is also a key tool for many areas of science such as quantification of global-scale geodynamical phenomena or studies on the atmosphere.

Over the past years, modern smartphones facilitate the collection of raw GNSS measurements. These affordable smart devices are utilized by billions of people making crowdsourcing of GNSS observations possible. This forms a great opportunity for investigating the usage of smartphone-based GNSS data for different subjects such as weather forecasting or atmosphere studies at large. By combining the observations acquired by high-quality GNSS stations with a dynamic network of GNSS-capable Internet-of-Things (IoT) devices, the sample of observations can be significantly increased and an unprecedented spatio-temporal resolution can be achieved.

2. PROJECT OVERVIEW

Application of Machine Learning Technology for GNSS IoT data fusion (CAMALIOT) is an ongoing project funded by the European Space Agency (ESA) that aims to collect large amounts of GNSS observations by developing an Android application and conducting a dedicated crowdsourcing campaign.

3. CROWDSOURCING CAMPAIGN

Campaign to collect data from satellites for scientific research

- Started on 17th March 2022
- Android app with 35k+ installations
- 10k registered users
- 10 billion GPS observations collected

4. SCIENCE USE CASE

ML-based forecasting of tropospheric parameters

One science use case is to utilize the IoT GNSS data to investigate its feasibility for augmenting modelling and prediction of tropospheric parameters such as Zenith Wet Delay (ZWD). ZWD is strongly related to the water vapour content in the atmosphere, which in turn drives weather systems and climate change to a great extent.

The aim is to derive a global ML-based model enhanced by Earth Observation data that can predict ZWD in space and time.

DATA

TARGET
Zenith Wet Delay
source: Nevada Geodetic Laboratory

FEATURES
Meteorological Data, Latitude, Longitude, Height, Time
source: ECMWF ERA5

TRAINING DATA - 80%
ZWD (y_train) features (X_train)

TESTING DATA - 20%
ZWD (y_test) features (X_test)

ML ALGORITHM
XGBoost
- standardize features
- hyperparameter optimization, cross-validation
- train model

EVALUATION
- make ZWD predictions for testing data (y_test_pred)
- calculate performance metrics
  - RMSE = 12.5 mm
  - R^2 = 0.97

output: ML model

apply ML model globally

5. SUMMARY

- CAMALIOT will advance current knowledge and experience concerning GNSS observations as collected from smartphones and applied for scientific purposes.
- Global ML-based ZWD model can make predictions for every point on Earth (assuming meteorological data is available) and achieves an average RMSE of 12.5 mm

Fig. 1: A high-level overview of the concept exploited within the CAMALIOT project.

Fig. 2: Map showing the number of files uploaded from different countries (until 29th April 2022).

Fig. 3: ML-based global ZWD predictions for the 5th July 2019, 00:00 UTC.