AUTH

UoP

# Modelling received power from wireless networks in Greece using machine learning


**by**

**Prof. George Tsoulos[a], Prof. Sotirios Goudos[b]**

[a]Dept. of Informatics & Telecommunications,
University of Peloponnese, Tripoli, Greece
gtsoulos@uop.gr

[b]School of Physics
Aristotle University of Thessaloniki, Thessaloniki
Greece
sgoudo@physics.auth.gr


WMCL

# Presentation Outline

- Measurement Campaign
- Machine Learning Methods basics
- Problem Definition
- Results
- Conclusions

# Motivation and objectives

- To predict RSRP from Base Stations based on Machine Learning for **4G and 5G Networks**
- To reduce prediction error for extreme cases
- To provide a prediction framework for similar cases

AUTH UoP

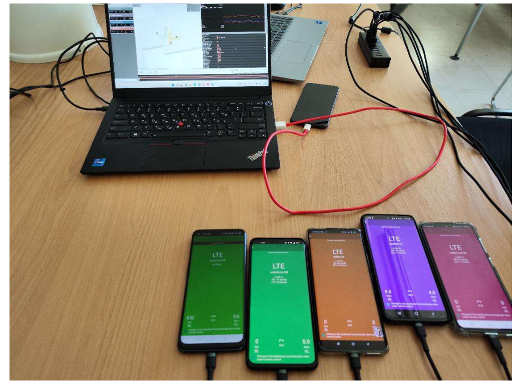
# Wireless and Mobile Communications Lab

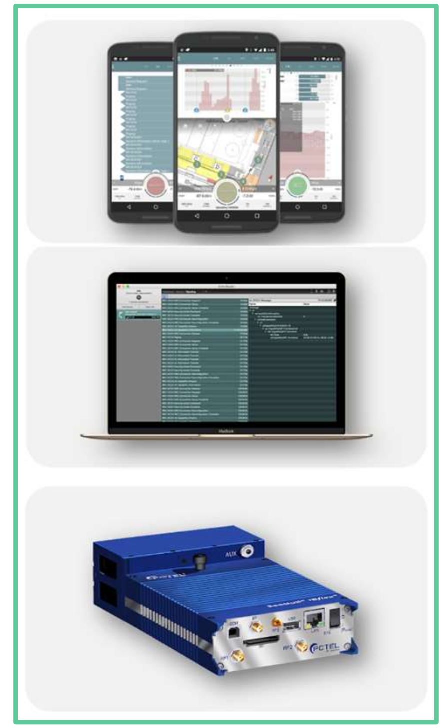
ACADEMIC RESEARCH WITH INDUSTRIAL APPLICATION


- Established in 2002; Based on more than 30 years of experience in R&D projects
- Lab members: **G. Tsoulos (Director)**, G. Athanasiadou, A. Kaloxylos, D. Zarbouti, D. Kontaxis, more than 10 postgraduates
- Equipped with state-of-the-art wireless network and system measurement tools (9 test mobiles, 2 scanners), EMF testing equipment (Narda SRM3006, AMB8059, AMS8061), and multiple drones (4)
- Current focus on modern wireless technologies, propagation – EMF - system measurements, and the development of smart wireless networks that utilize multiple sensors and drones

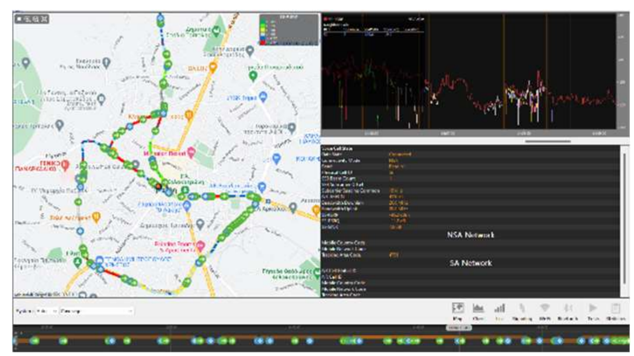
# Wireless Systems Measurements

- Echo Suite from Enhancell
  - Echo One
  - Echo Studio
  - Echo Cloud





- **Analyzes** the radio link and captures systemic parameters such Beams, Cell ID, Channel Frequency, Connectivity Mode, Modulation, RSRP, RSRQ, SINR, neighbor RSRP, Latency, Throughput, etc.



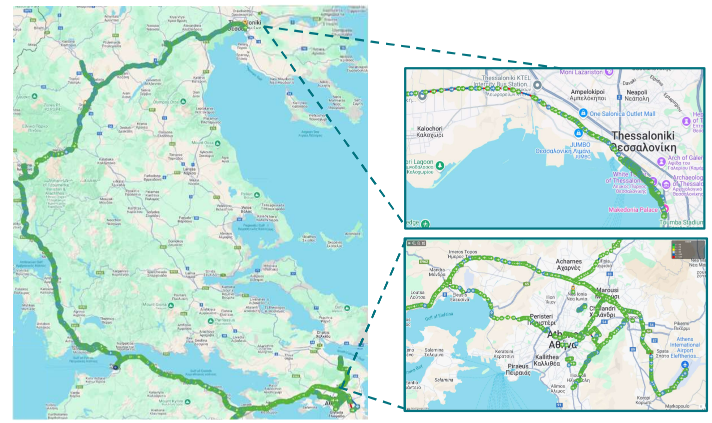
- Simultaneous measurements for the three wireless networks in Greece (Cosmote, Vodafone, Nova)

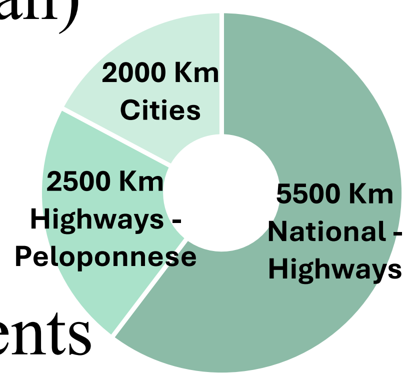
# Measurement Campaign (1)

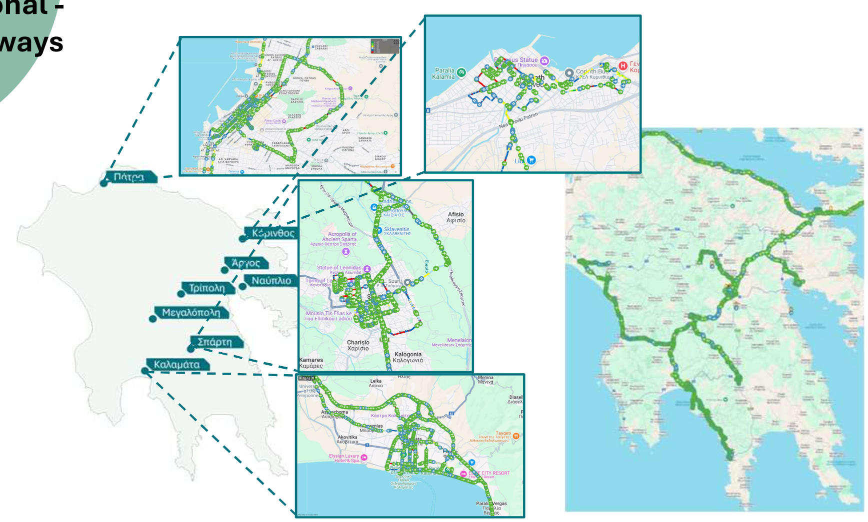
**Vehicle based data collection across different operators**

- 3 smartphones on the car's dashboard

- Routes
  - cities (urban-suburban)
  - Highways (rural)





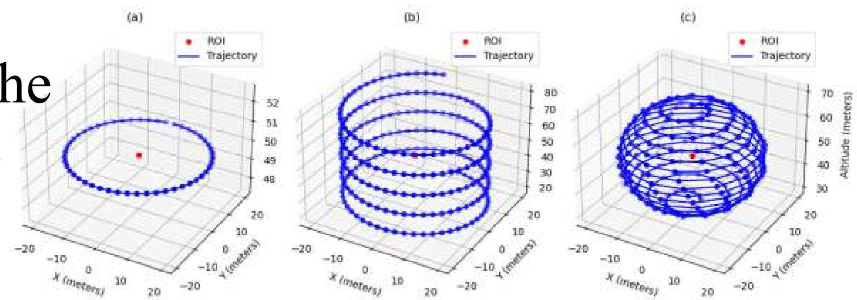




- Continuous measurements for signal/network/service quality (power, interference, throughput, latency, etc.)

# Measurement Campaign (2)

**Drone-assisted data collection across different operators**

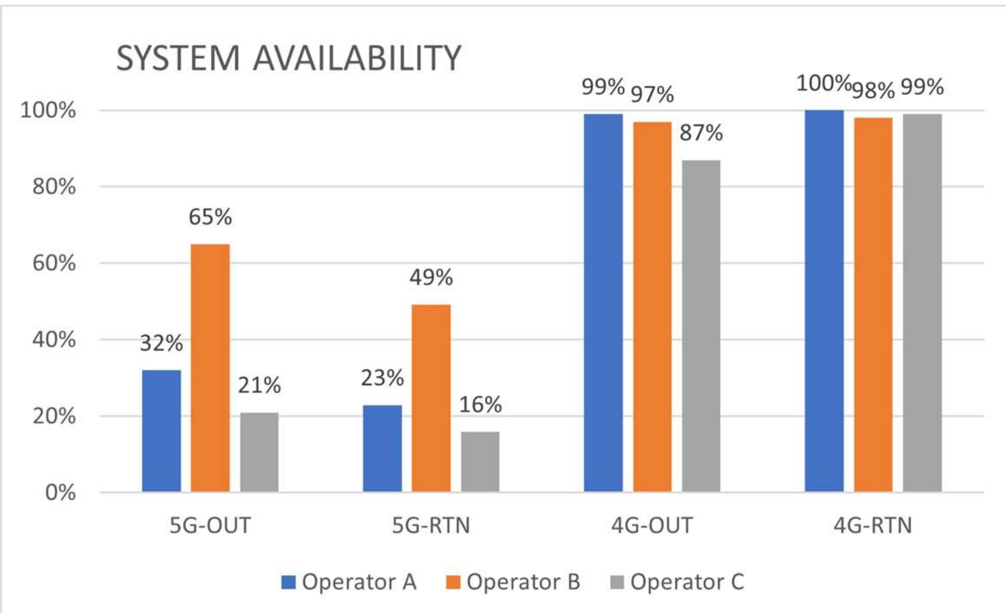
- Industrial Hexacopters with 3 smartphones
- Max payload 10 kg, max speed 15 m/s, max flight duration 75 min, max range 10 km
- Ground station monitored and controlled the experimental process. Mission Planner for planning, configuring, simulating, and monitoring autonomous missions.
- Continuous measurements for signal/network/service quality (power, interference, throughput, latency, etc.)



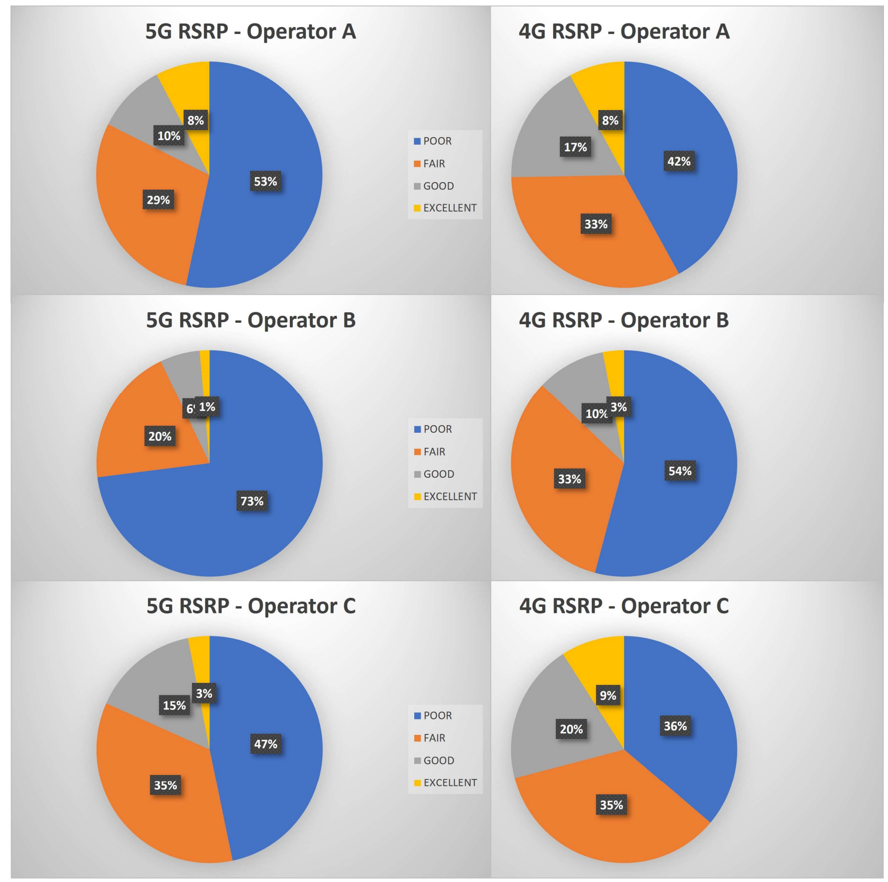

AUTH UoP

# Field Trials (1)

- **SYSTEM AVAILABILITY**




- **Coverage Quality**



# Field Trials (2)

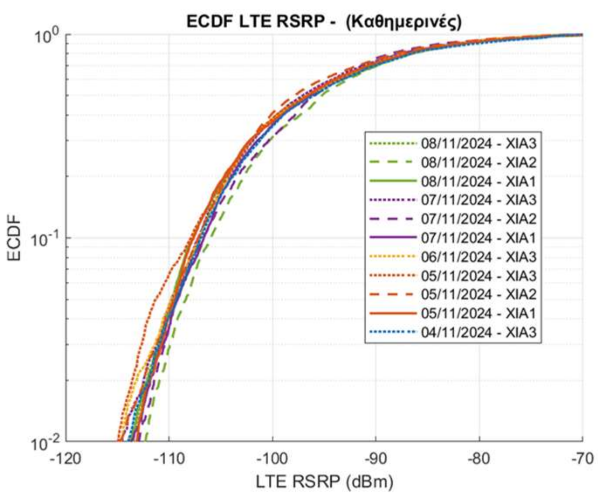
- **4G-5G RSRP COLOR MAPS**

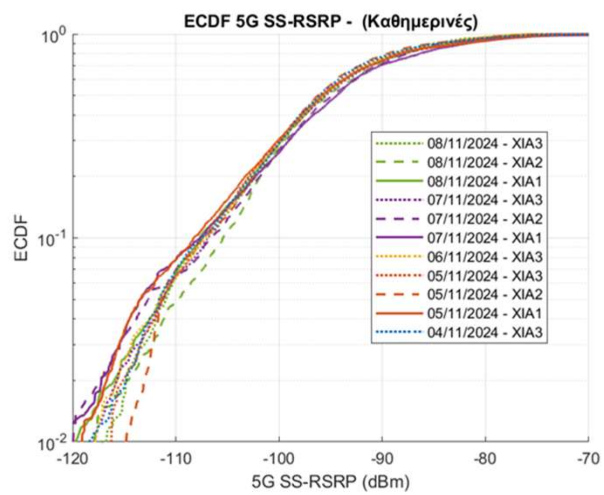
AUTH

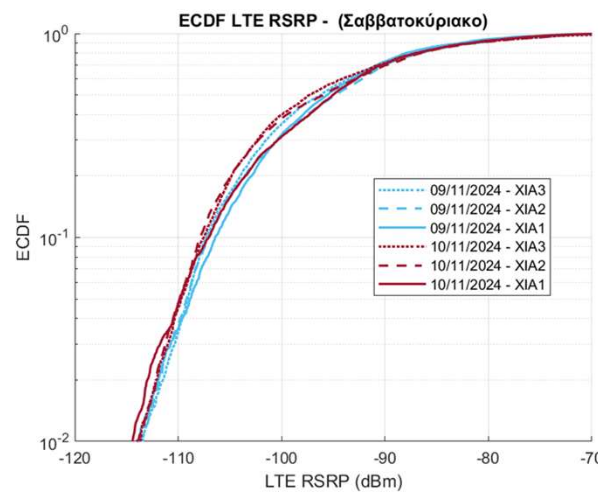
UoP

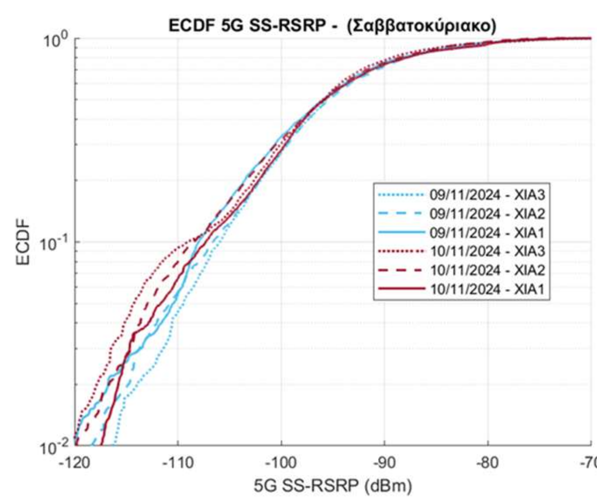
# Field Trials (3)

- **4G-5G RSRP statistics – different days/time of day**



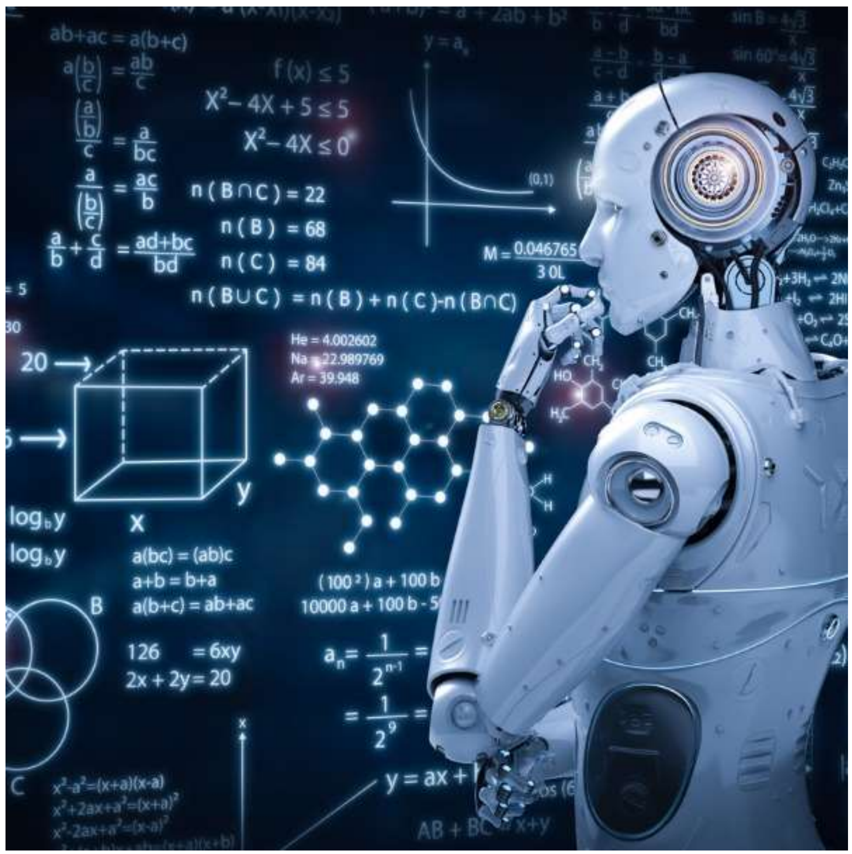



AUTH UoP

# Machine Learning Methods



AUTH UoP

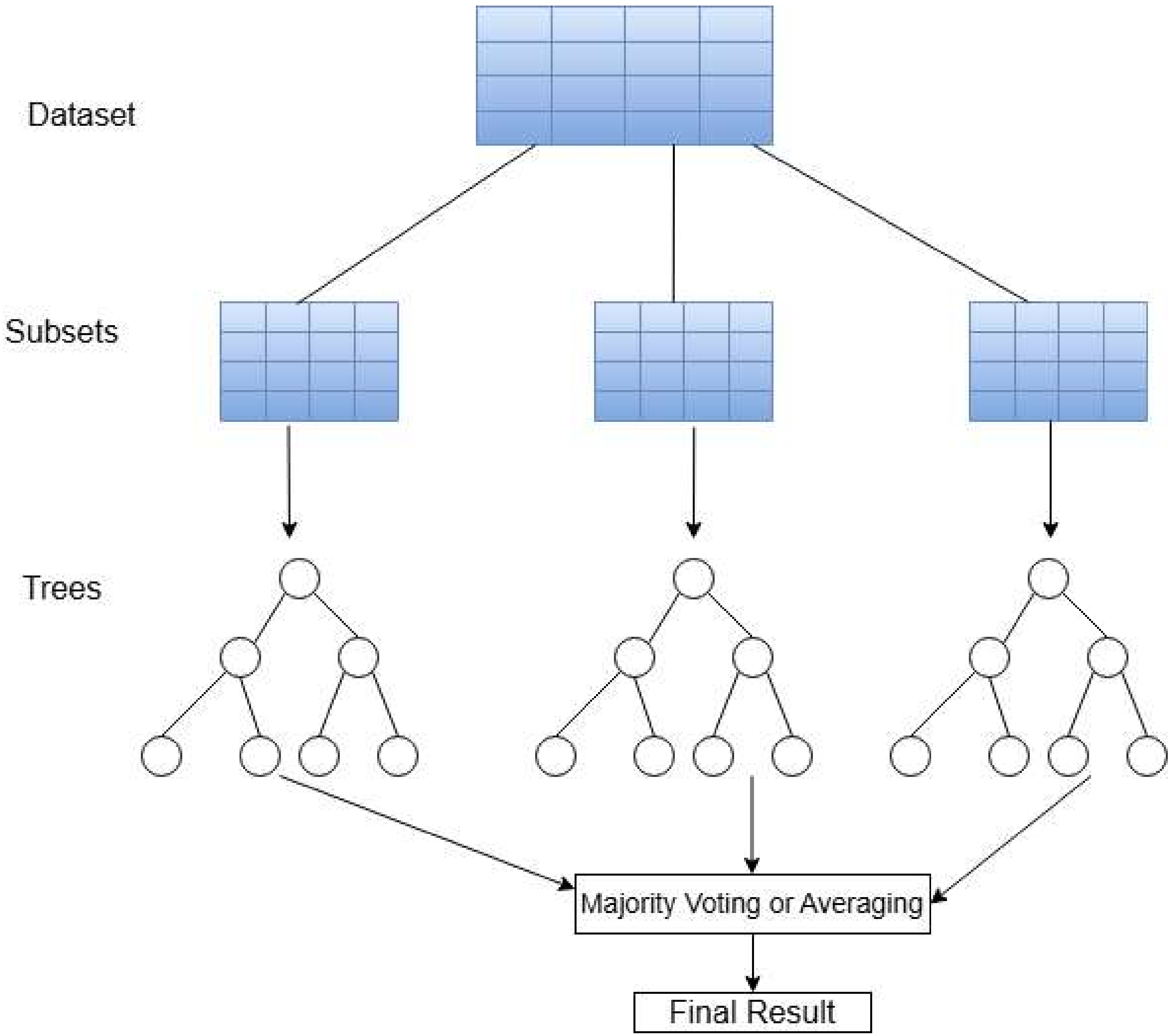
# ELEDIA@AUTH, Department of Physics, AUTH
# Director:
# Prof. Sotirios K. Goudos

ELEDIA@AUTH is a Research Group in the AUTH, Department of Physics consisting of 5 faculty members, 5 PhD students, and several MSc students

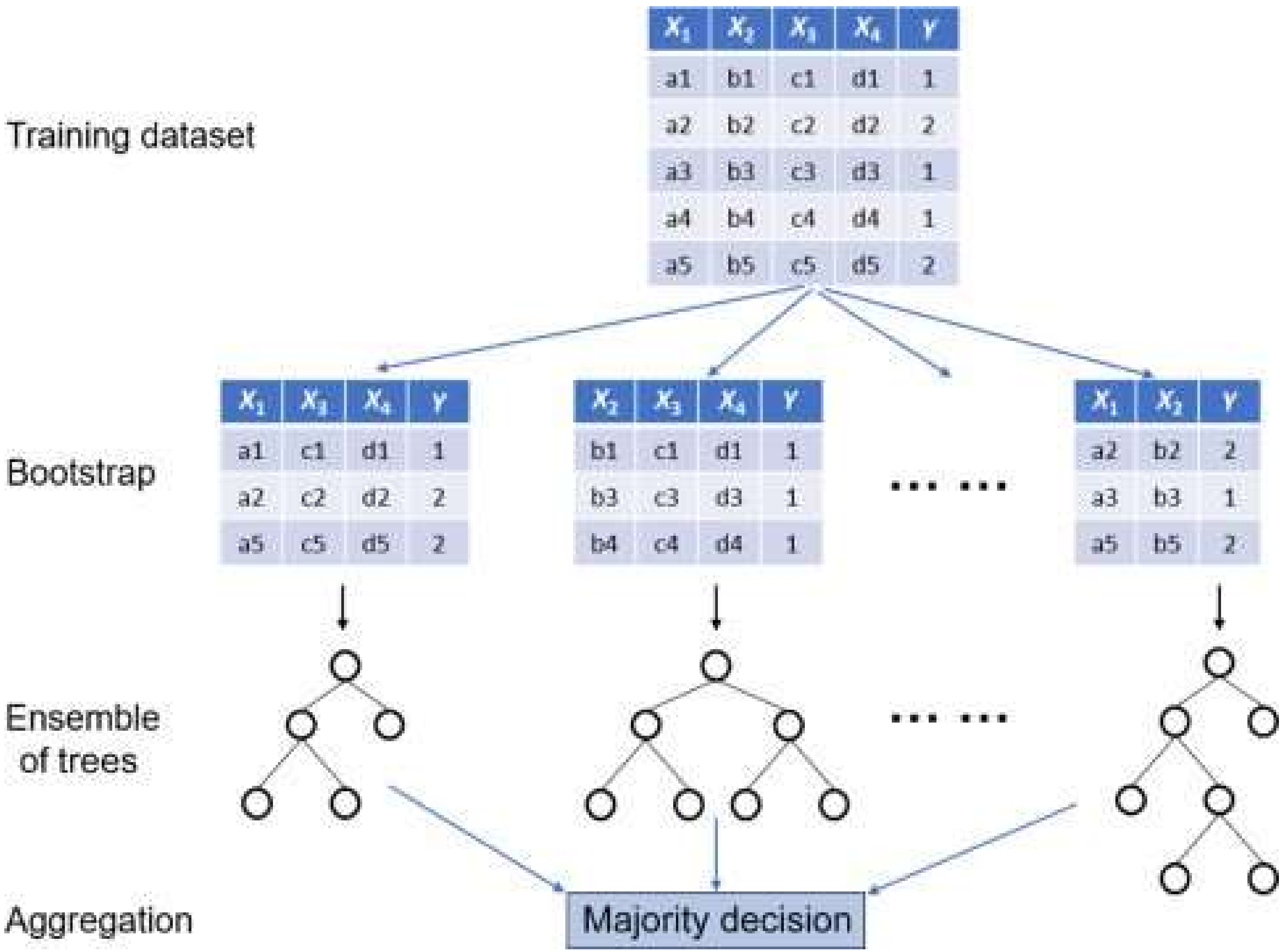
# ML Methods used

- Bagging methods
  - Random Forest, Extra trees
- Boosting methods
  - CatBoost, XGBoost, LGBM
- ANN

# ML Methods: Random Forest (2)

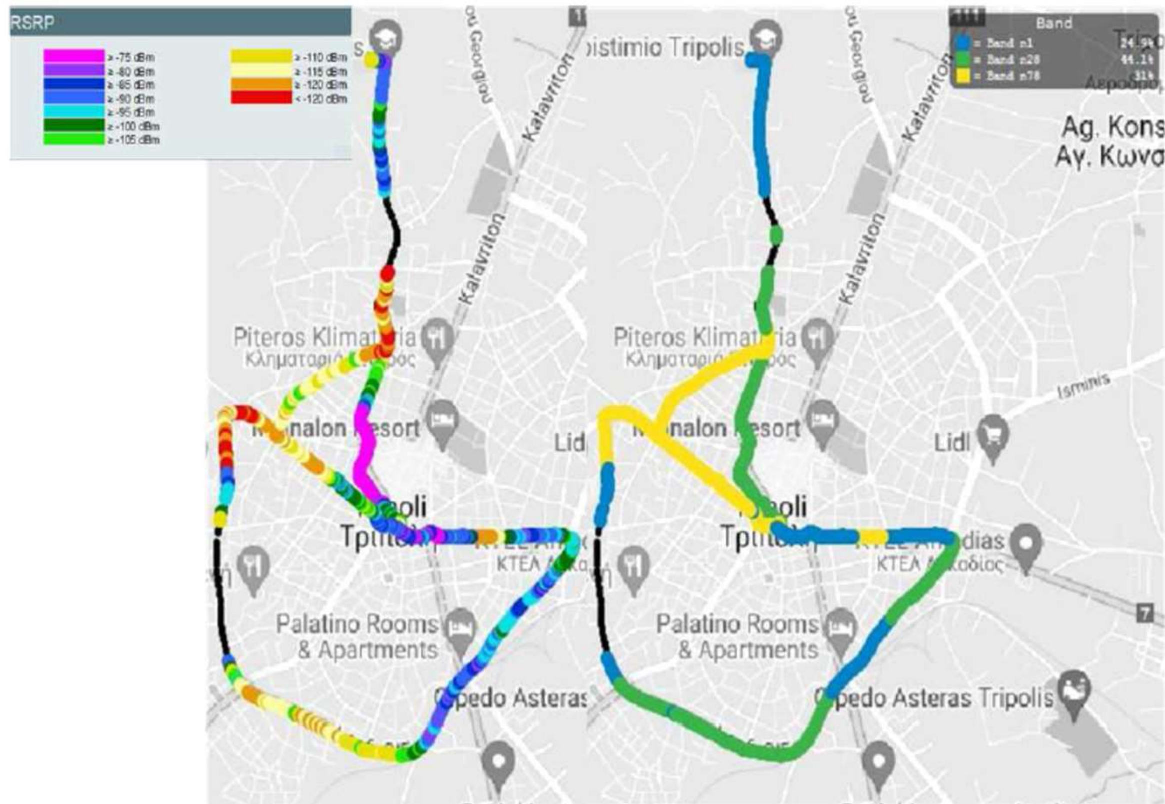


# ML Methods: Extremely Randomized Trees



AUTH UoP

# Problem Definition

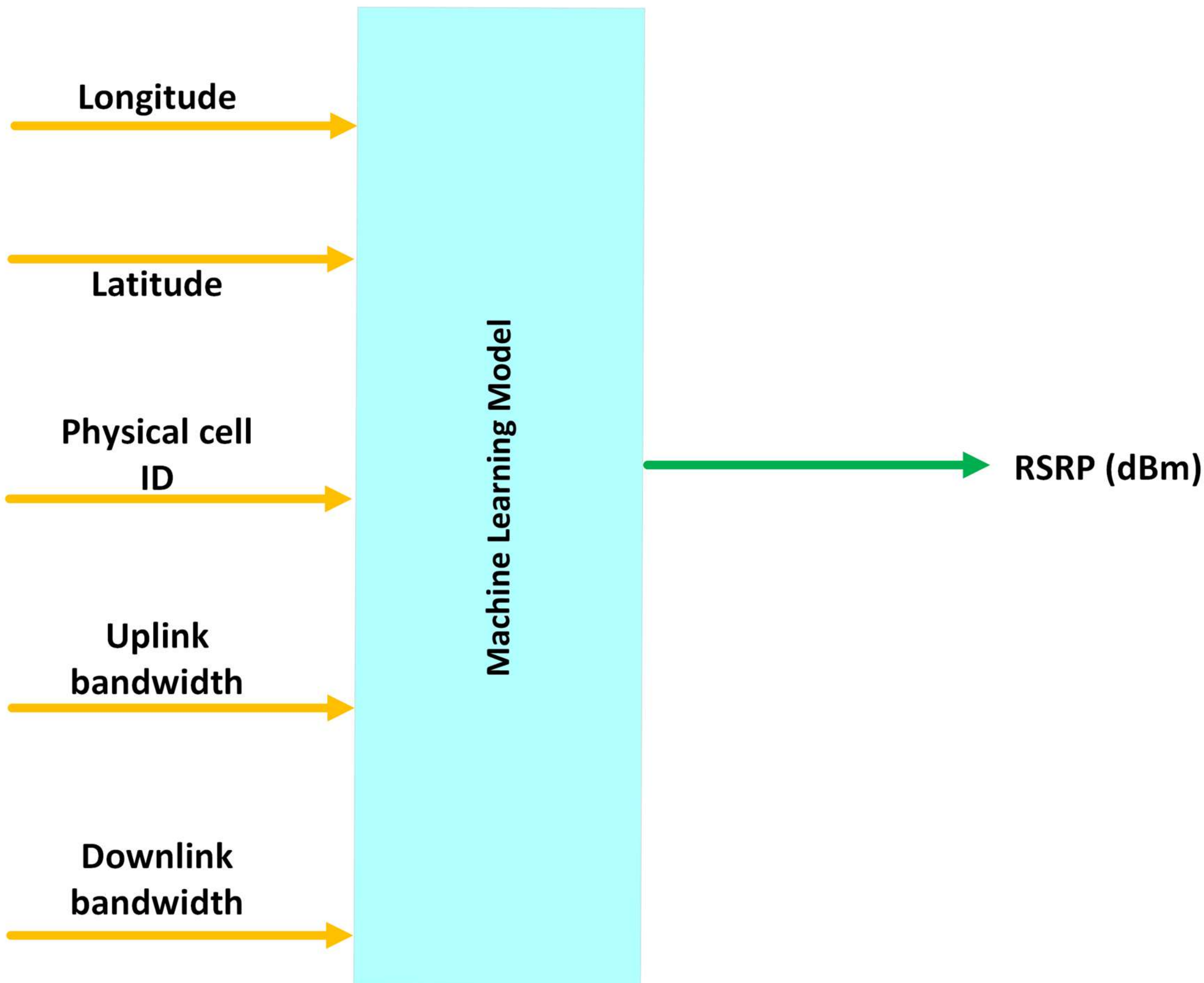



Goal: to use **ML** for **SS-RSRP prediction** (linear average of power contributions (in [W]) of resource elements that carry secondary synchronization signals )

# Problem Definition (2)

- Tripoli, a city in southern Greece **12Km** route
- Enhancell's Echo One
- UE device assesses all cellular networks (2G-5G), examining various parameters
- Two datasets are constructed based on the 4G (5062 vectors) and 5G data (3236 vectors)
- **Optuna for hyperparameter optimization**
- 80% - 20% training-test split

# Machine Learning Model



# Performance metrics for Regression

Real RSRP

Predicted RSRP

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \overline{y}_i|$$

$$RMSE = \sqrt{(\frac{1}{n})\sum_{i=1}^{n}(y_i - \overline{y}_i)^2}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \overline{y}_i}{y_i}\right| x100\%$$

$$\mathrm{SMAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\frac{|\overline{y}_i - y_i|}{(|\overline{y}_i| + |y_i|)/2}$$

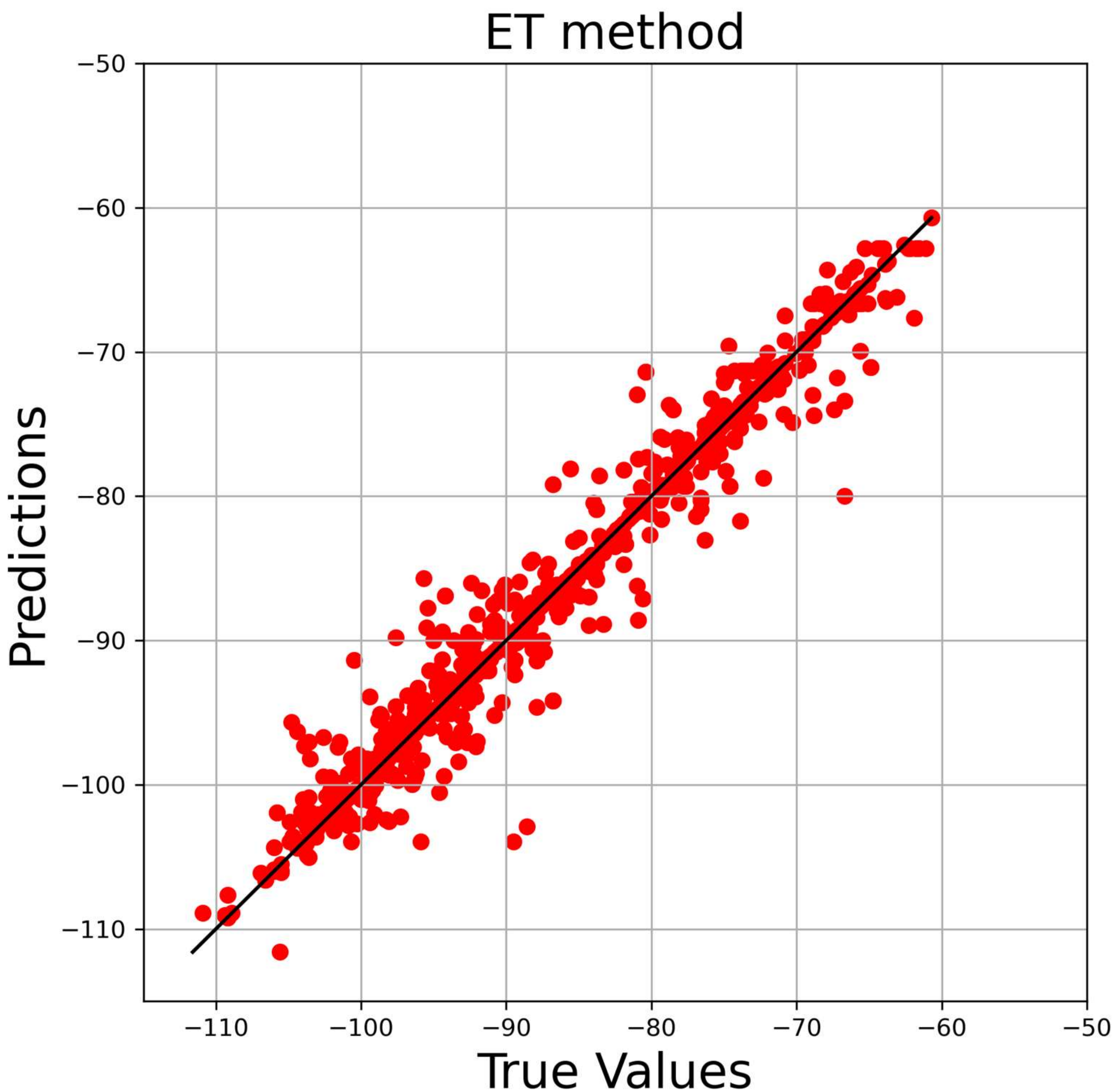
# Numerical results 4G

| ML Model | MAE (dBm) | RMSE (dBm) | MAPE (%) | SMAPE (%) | Training Time (s) |
|---|---|---|---|---|---|
| XGBoost | 1.32 | 2.21 | 1.53 | 0.76 | 119.22 |
| LightGBM | 1.32 | 2.41 | 1.7 | 0.87 | 1.51 |
| CatBoost | 1.43 | 2.38 | 1.64 | 0.83 | 124 |
| Random Forest | 1.28 | 2.13 | 1.51 | 0.75 | 25.17 |
| **Extra Trees** | **1.18** | **2.16** | **1.39** | **0.69** | **2.78** |
| ANN | 1.87 | 2.97 | 2.38 | 1.19 | 105.19 |

AUTH

# Numerical results 4G

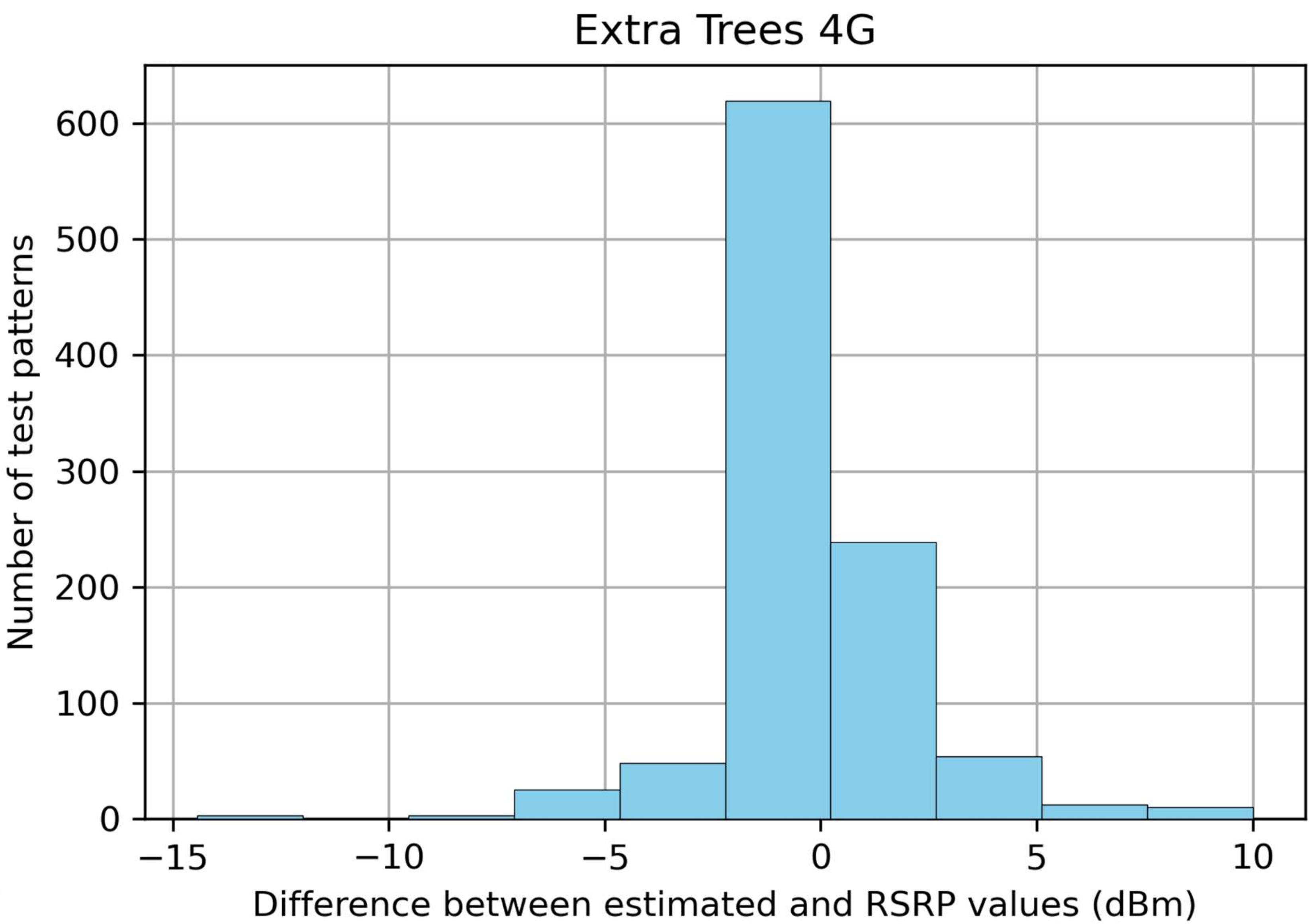
UoP



AUTH

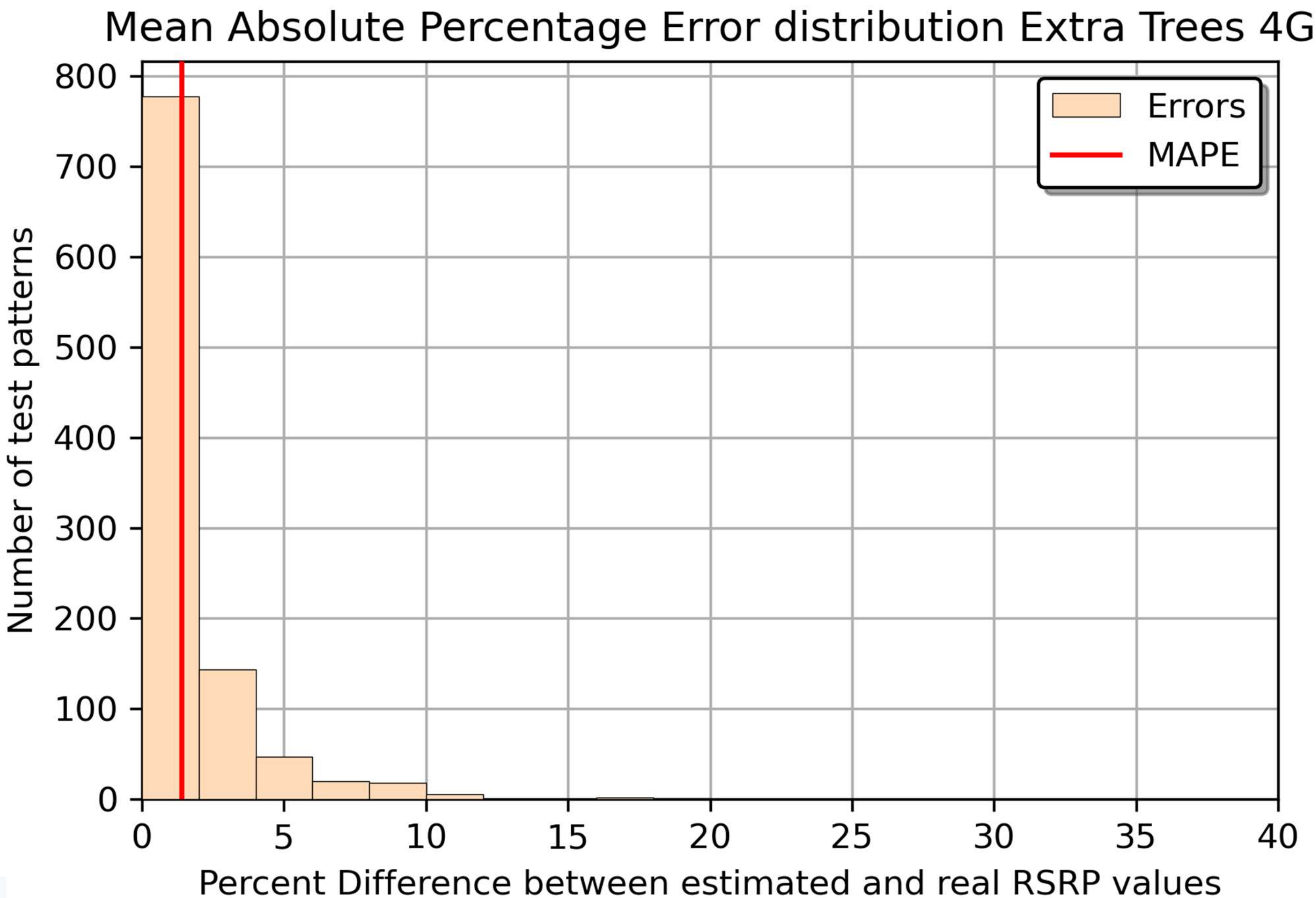
# Numerical results 4G

UoP

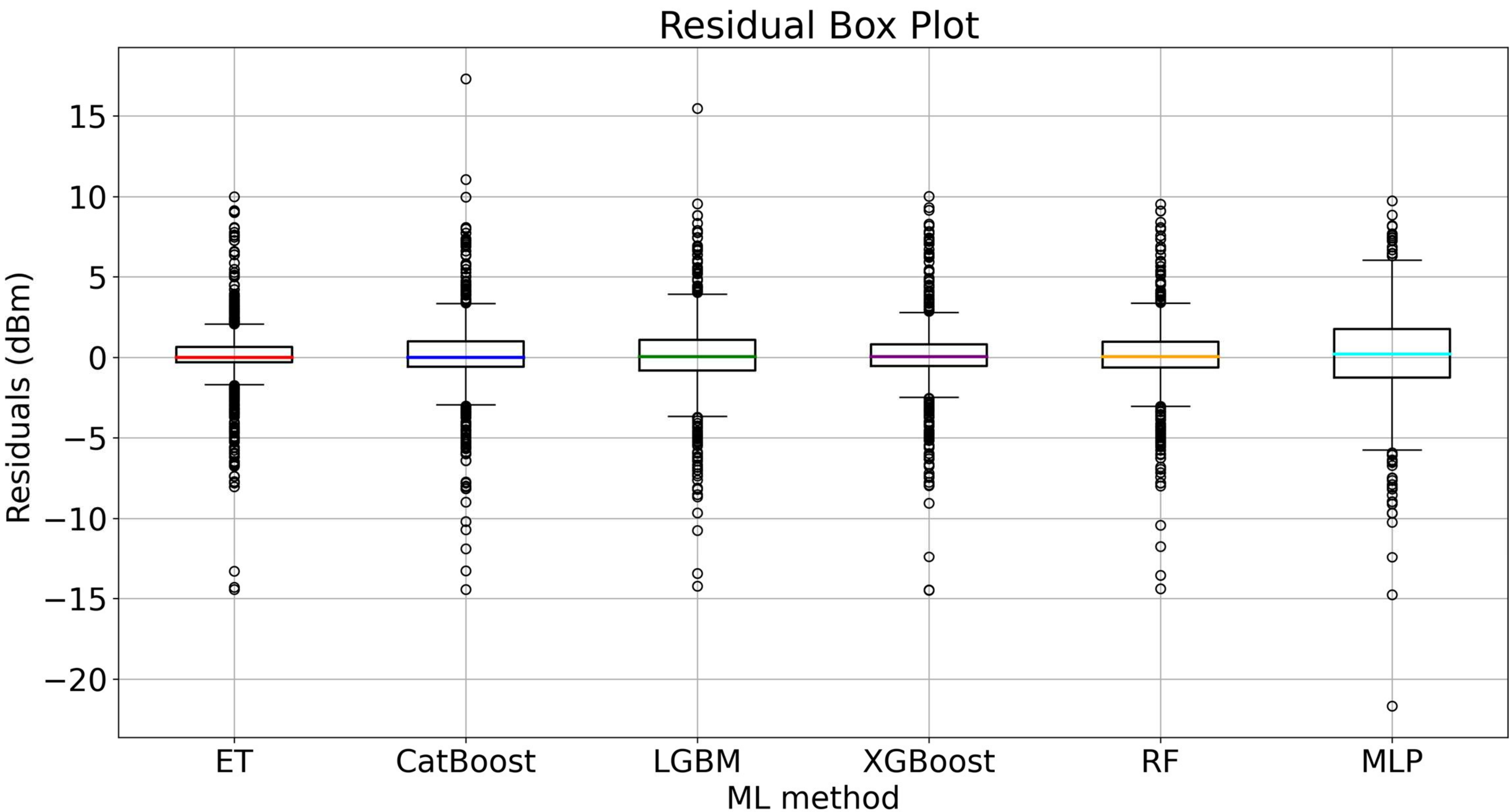


AUTH UoP

# Numerical results 4G



# Numerical results 4G



AUTH

UoP

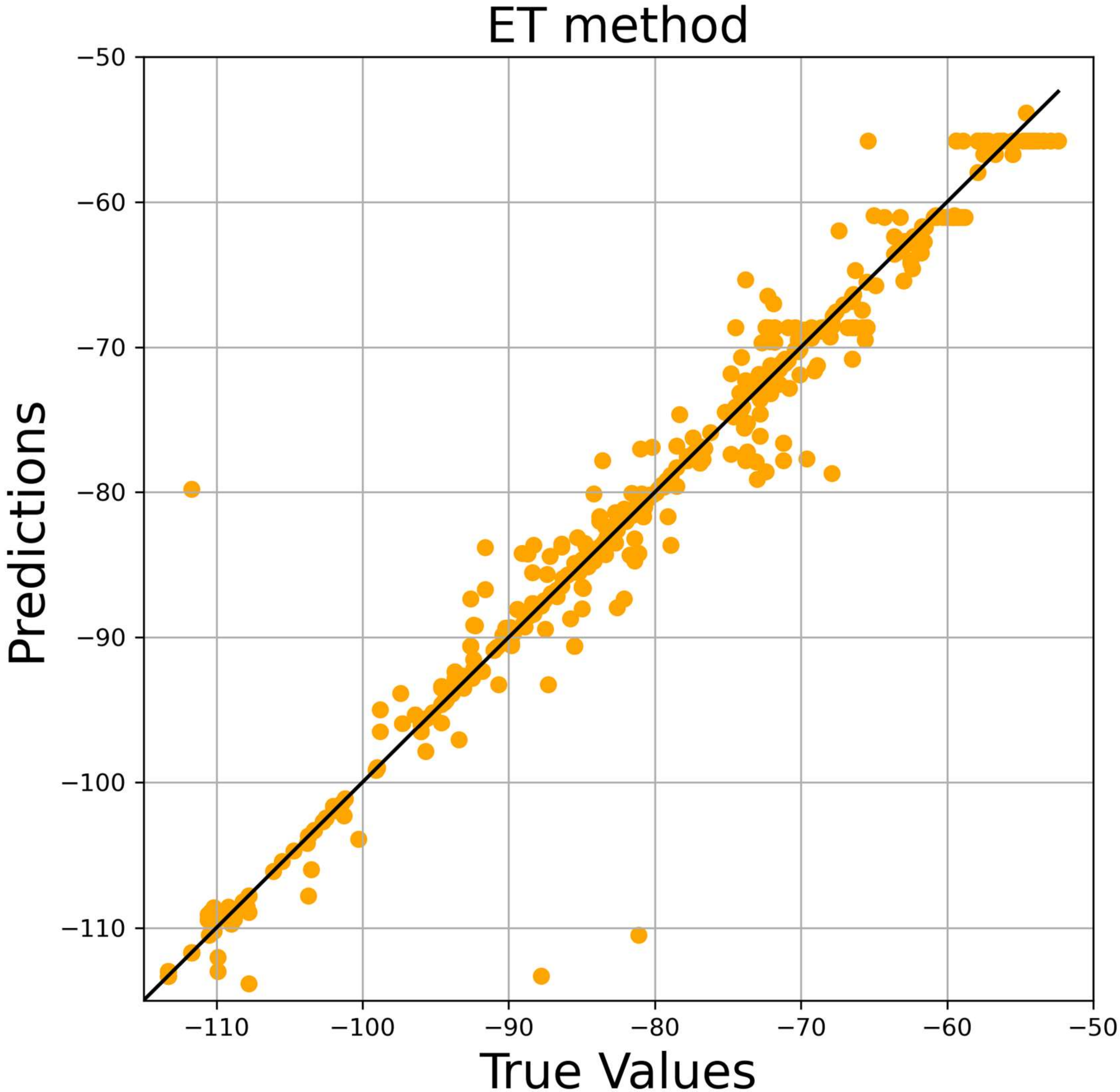
# Numerical results 5G

| ML Model | MAE (dBm) | RMSE (dBm) | MAPE (%) | SMAPE (%) | Training Time (s) |
|---|---|---|---|---|---|
| XGBoost | 1.29 | 2.7 | 1.7 | 0.84 | 74.12 |
| LightGBM | 1.46 | 2.41 | 1.69 | 0.9 | 3.47 |
| CatBoost | 1.52 | 2.33 | 1.79 | 0.89 | 134.24 |
| Random Forest | 1.17 | 2.64 | 1.55 | 0.76 | 3.83 |
| **Extra Trees** | **1.29** | **2.18** | **1.43** | **0.71** | **2.72** |
| ANN | 1.94 | 3.01 | 2.46 | 1.23 | 101.38 |

AUTH

# Numerical results 5G

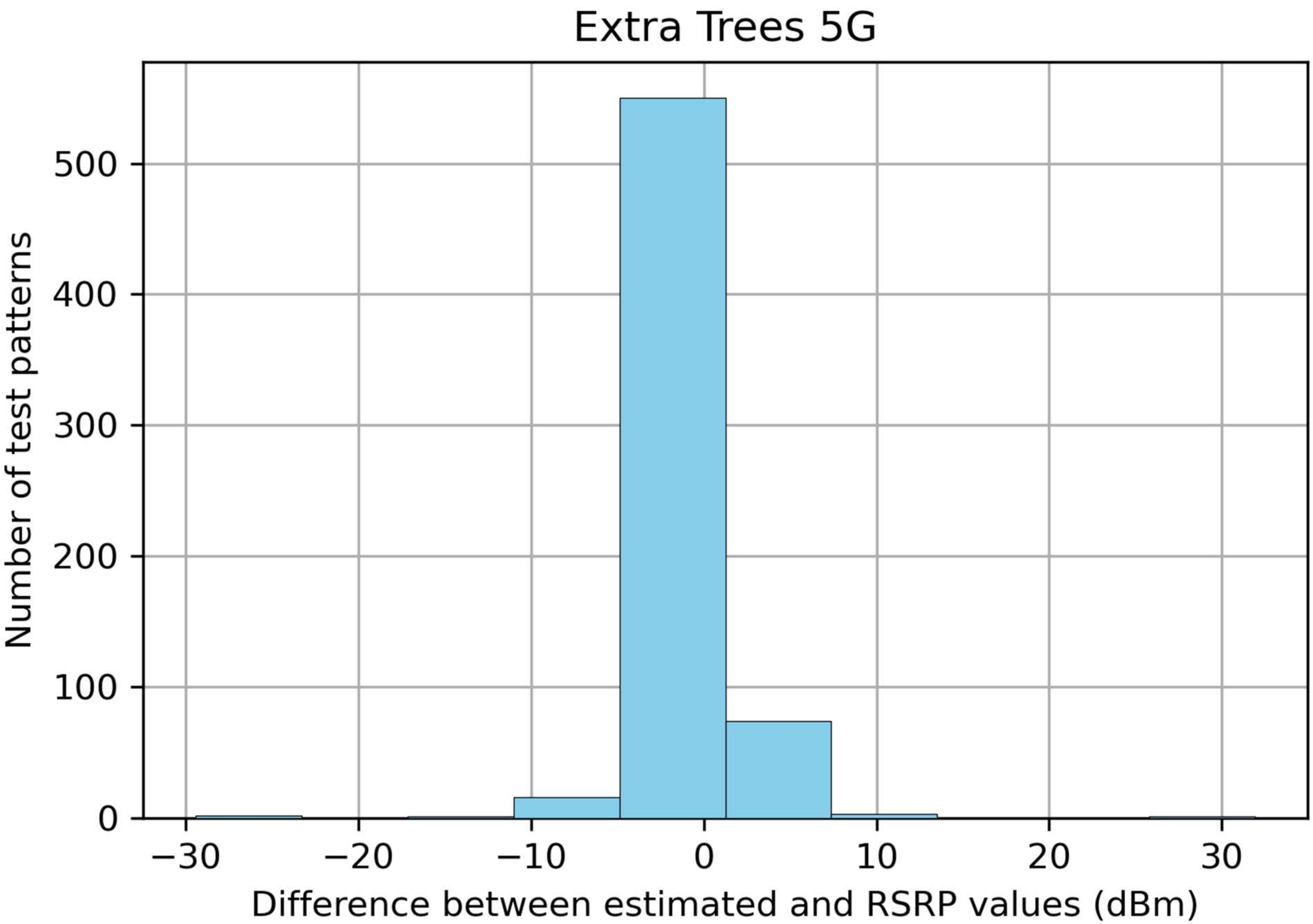
UoP



AUTH

# Numerical results 5G

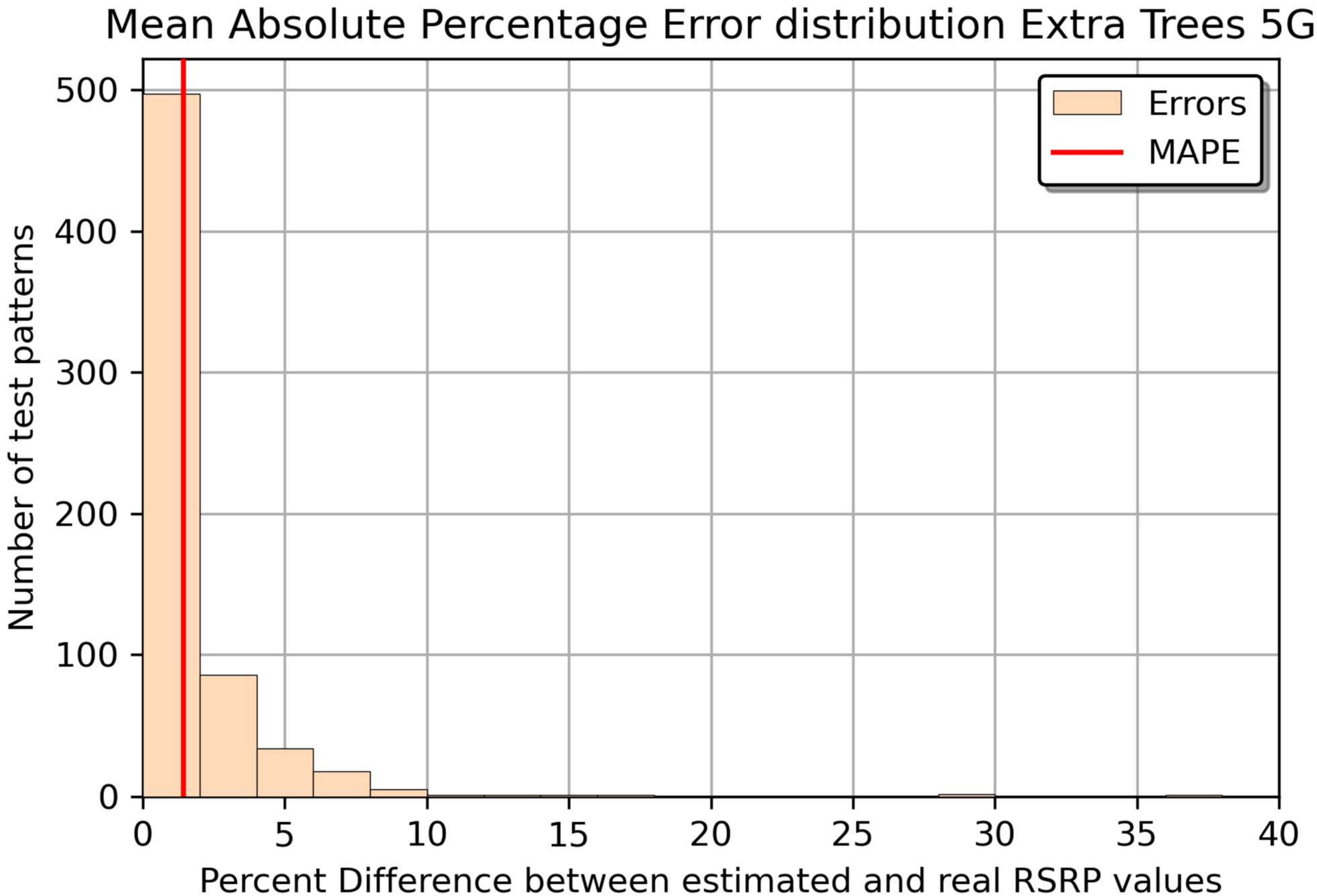
UoP



AUTH

# Numerical results 5G

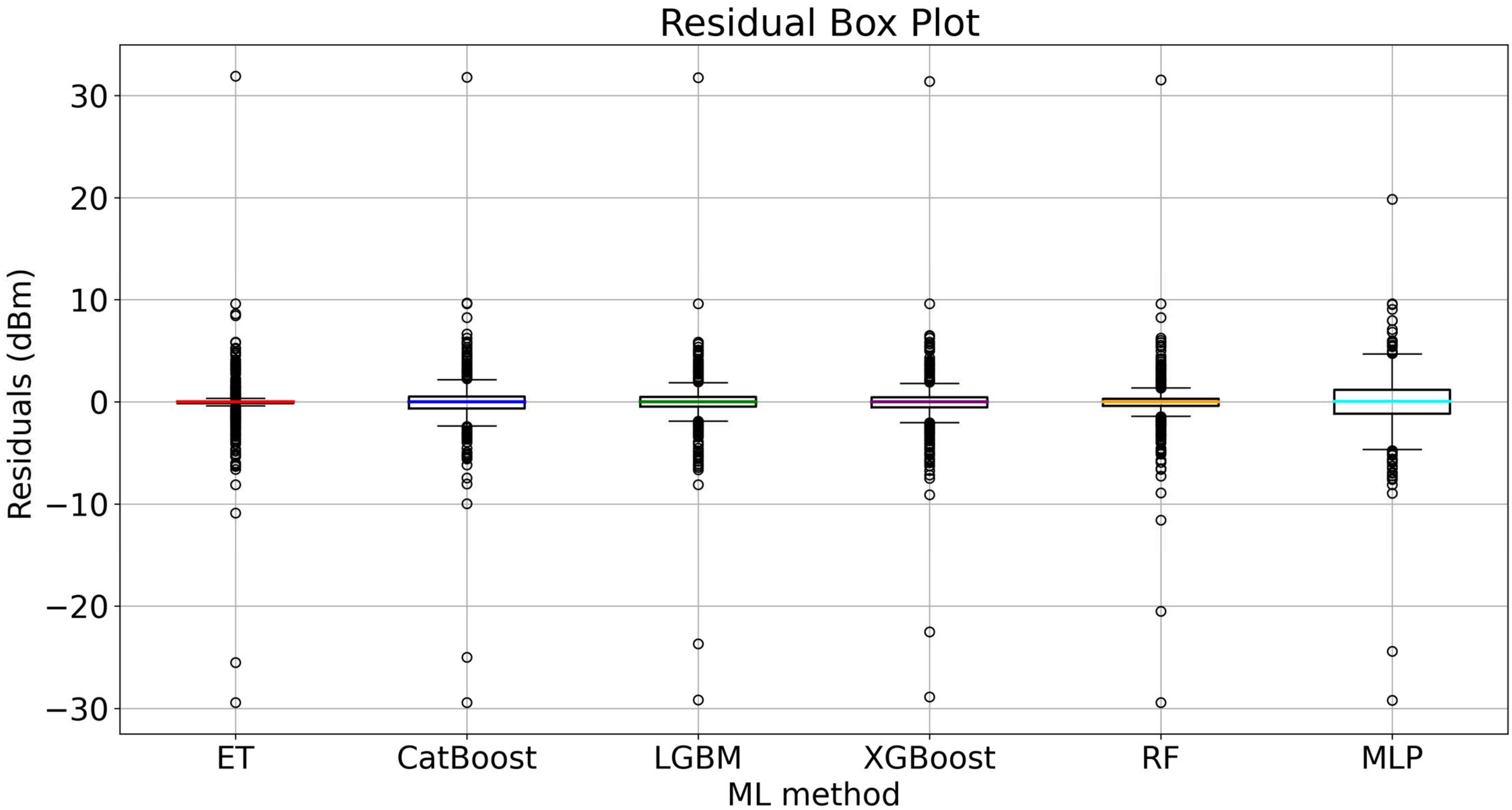
UoP



AUTH

# Numerical results 5G

UoP



# Conclusions

- The simulation results indicate significant accuracy improvements in estimation across various ML methods
- **Among these methods, ET performs best**
- ML approaches offer potential solutions for RSRP prediction challenges and can greatly assist future wireless network planning efforts in various environments for 4G and 5G scenarios

# Ensemble methods

- Basic idea of ensemble methods:
  - Combining predictions from competing models often gives better predictive accuracy than individual models.
- Shown to be empirically successful in wide variety of applications.
- Popular methods:
  - **Bagging: individual learners trained independently.**
  - **Boosting: training process is <u>sequential</u> and <u>iterative</u>**

# Bagging

- **Bagging = bootstrap + aggregation**

1. Create *k* bootstrap samples.

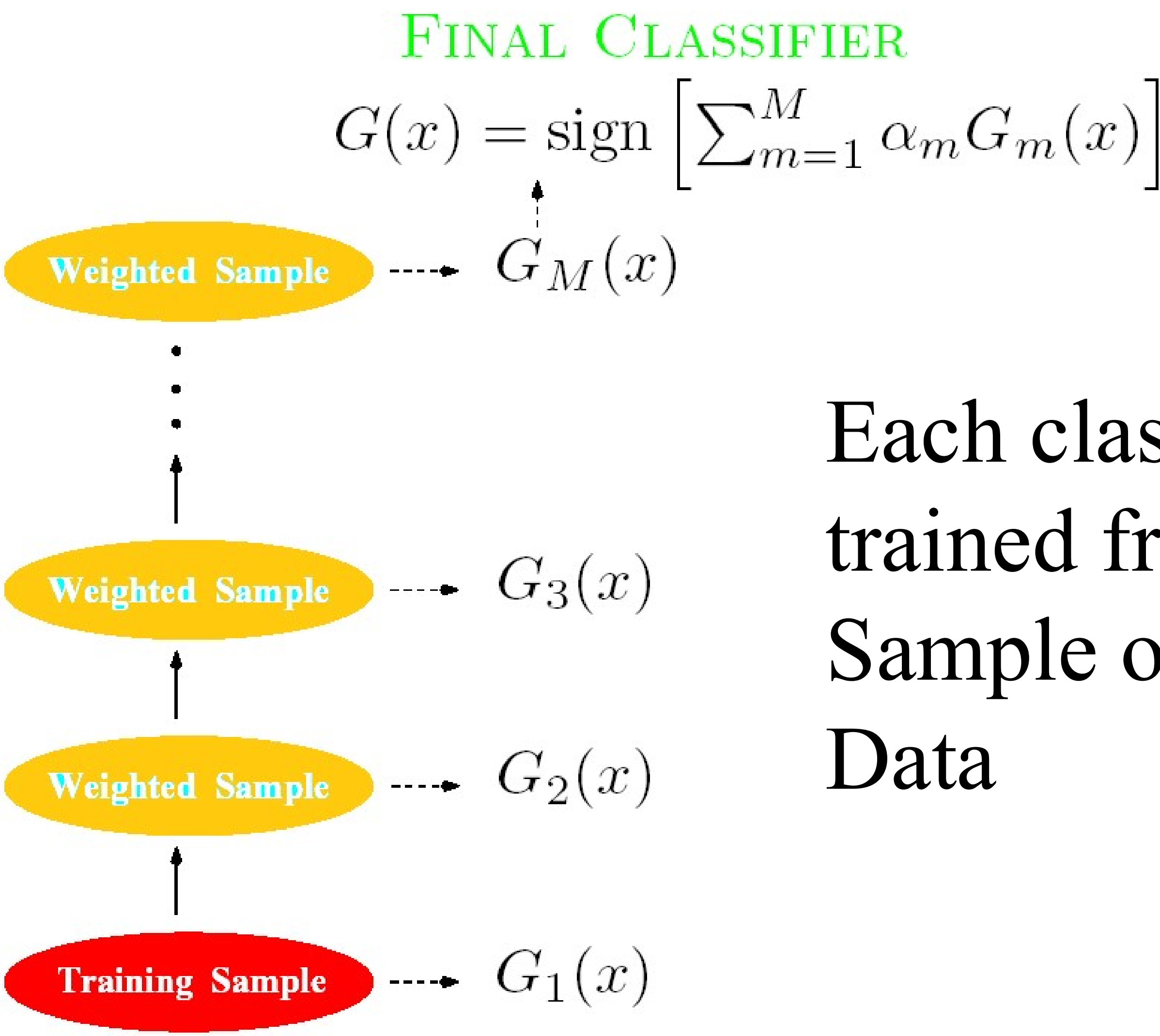
   Example:

| original data | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| bootstrap 1 | 7 | 8 | 10 | 8 | 2 | 5 | 10 | 10 | 5 | 9 |
| bootstrap 2 | 1 | 4 | 9 | 1 | 2 | 3 | 2 | 7 | 3 | 2 |
| bootstrap 3 | 1 | 8 | 5 | 10 | 5 | 5 | 9 | 6 | 3 | 7 |

2. Train a regressor on each bootstrap sample.
3. Average the predictions of the *k* models.

Breiman, L. "Bagging Predictors." *Machine Learning*. Vol. 26, pp. 123–140, 1996.




g

Each classifier $G_m(\mathbf{x})$ is trained from a weighted Sample of the training Data

# ML Methods: Categorical boosting (Catboost)

- A high-performance open-source gradient boosting library.
- Optimized for categorical data (no need for manual encoding).
- Built-in handling of missing values.
- Automatically converts categorical features to numerical ones

CatBoost

L. Prokhorenkova, et al,
CatBoost: unbiased boosting with categorical features
Proceedings of the 32nd International Conference on Neural Information Processing Systems (2018), pp. 6639-6649

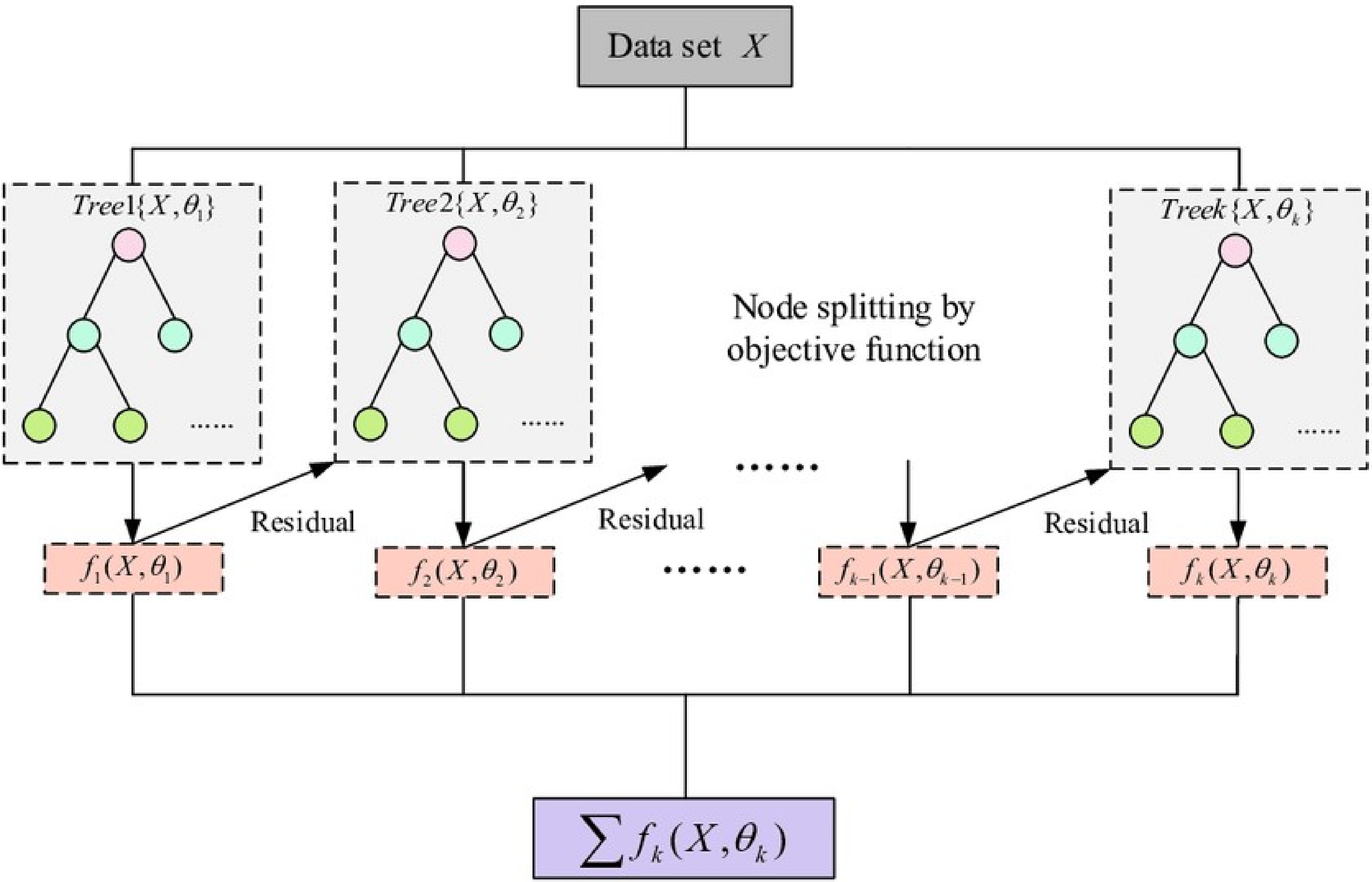
# ML Methods: Random Forest

- Based on Bootstrap Aggregation (Bagging)
- Each individual tree is grown indepedently from the others
- The final prediction is the average of the individual predictions
- L. Breiman, “Random forests,” *Machine Learning,* vol. 45, no. 1, pp. 5–32, Oct. 2001.
- P. Geurts, D. Ernst., and L. Wehenkel, “Extremely randomized trees”, Machine Learning, 63(1), 3-42, 2006.

# ML Methods: eXtreme Gradient Boosting (XGBoost)

- Based on Gradient Boosting
- Trees are grown sequentially
- Each tree compensates for the errors of the previous one
- T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* ser. KDD ’16. NewYork, NY, USA: ACM, 2016, pp. 785–794

# ML Methods: XGBoost (2)



# ML Methods: Light Gradient Boosting Machine (LightGBM)

Gradient boosting framework that uses decision trees

- ❑ Faster Training Speed and Efficiency
- ❑ High Accuracy
- ❑ Scalability
- ❑ Leaf-wise Tree Growth