R2DATO

A framework for GNSS-based solutions performance analysis in an ERTMS context

J. Marais

Université Gustave Eiffel

Q. Mayolle, M. Fasquelle

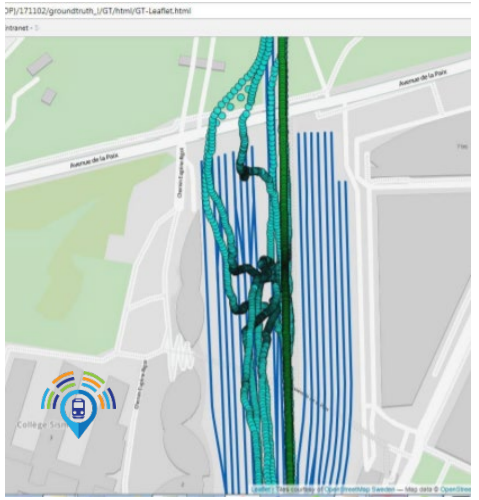
V. Tardif, E. Chéneau-Grehalle

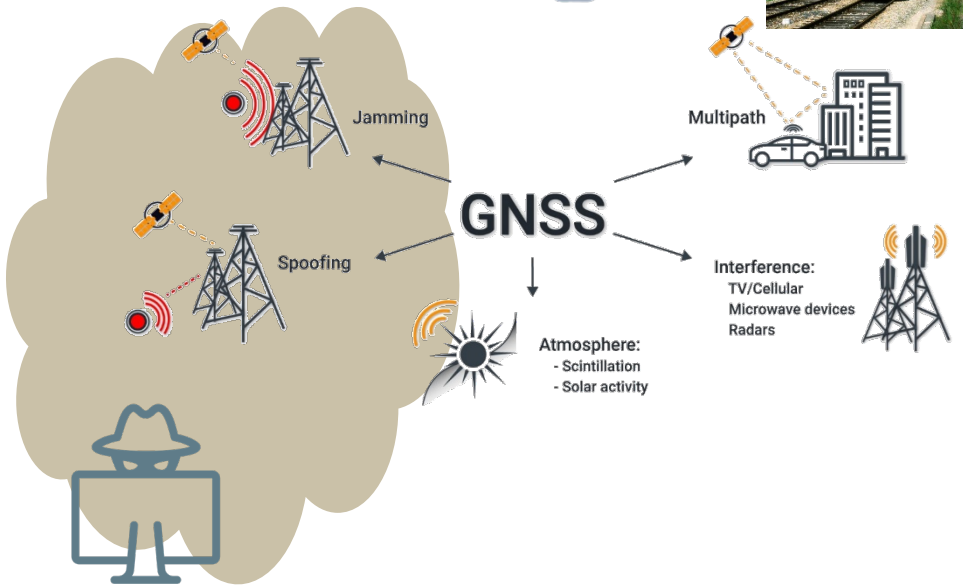
SNCF

RAILENIUM RAIL RESEARCH & INNOVATION

i-TRANS INNOVATION CATALYST

GNSS PERFORMANCE





Russia is jamming GPS satellite signals in Ukraine, US Space Force says

By Elizabeth Howell published 14 days ago

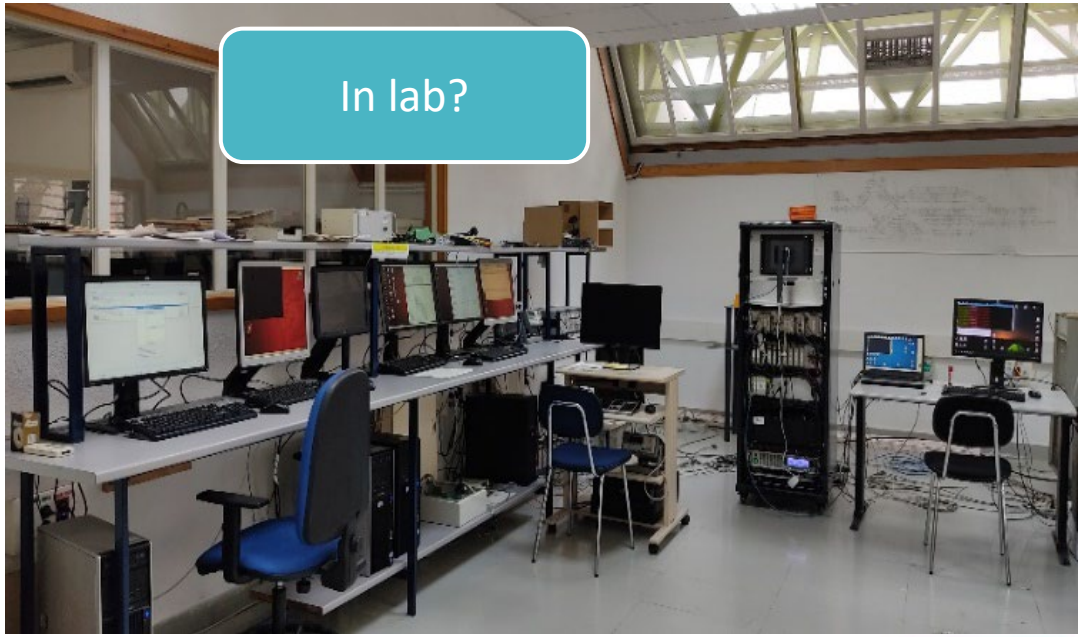
New GPS 'circle spoofing' moves ship locations thousands of miles

May 26, 2020 - By Dana Goward Est. reading time: 2 minutes

CONTEXT

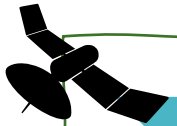
Progresses in GNSS-based solution introduction in rail applications – R2DATO

Solution development → Perf. evaluation → Safety demos → Certification

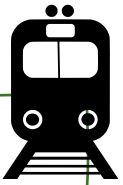


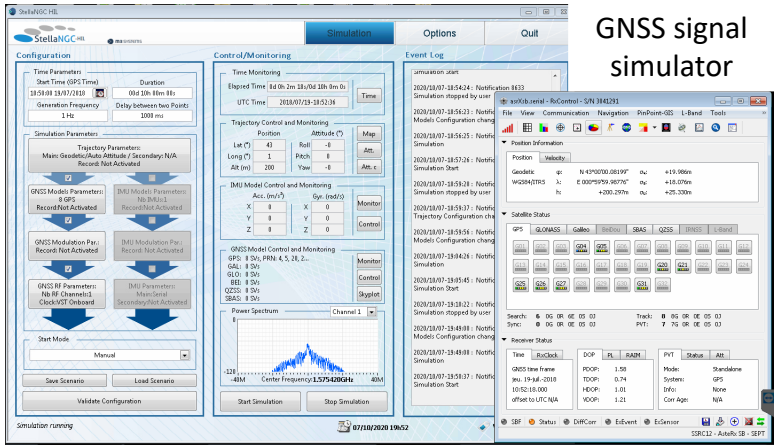
HOW TO SIMULATE GNSS ALONG A RAILWAY LINE?

The simulation chain linking space & rail



satellite signals → Propagation channel → Receiver processing → Integration in OB unit → Use in ERTMS

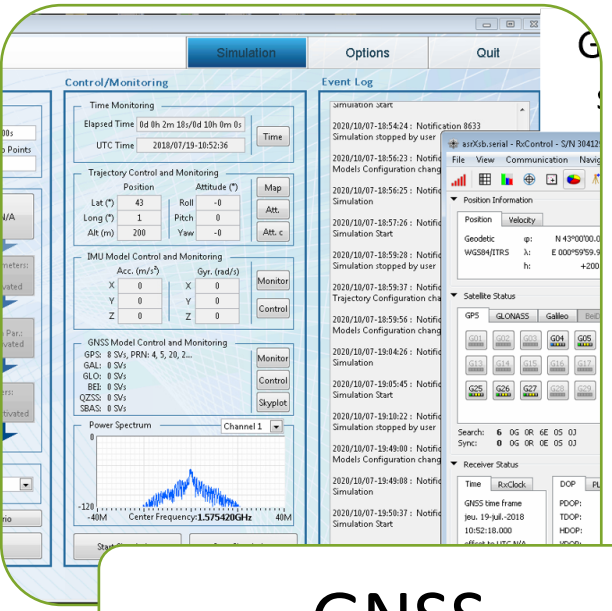




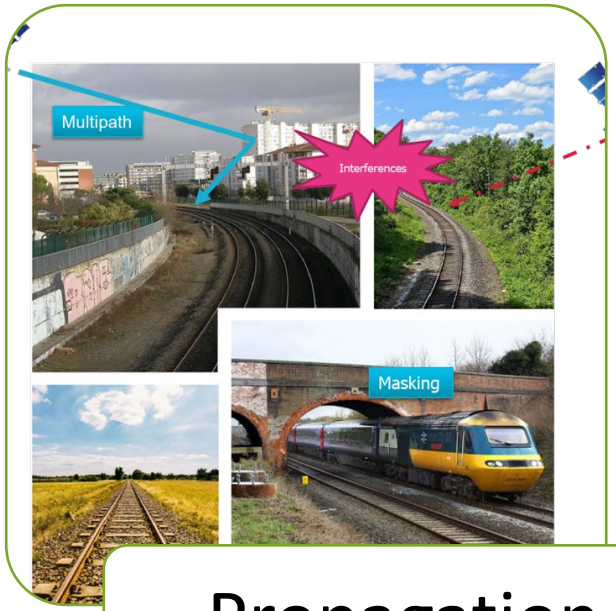
GNSS signal simulator

HOW TO SIMULATE GNSS ALONG A RAILWAY LINE?

The need: use of real(istic) railway errors



GNSS SIMULATION



Propagation channel

ERTMS simulation

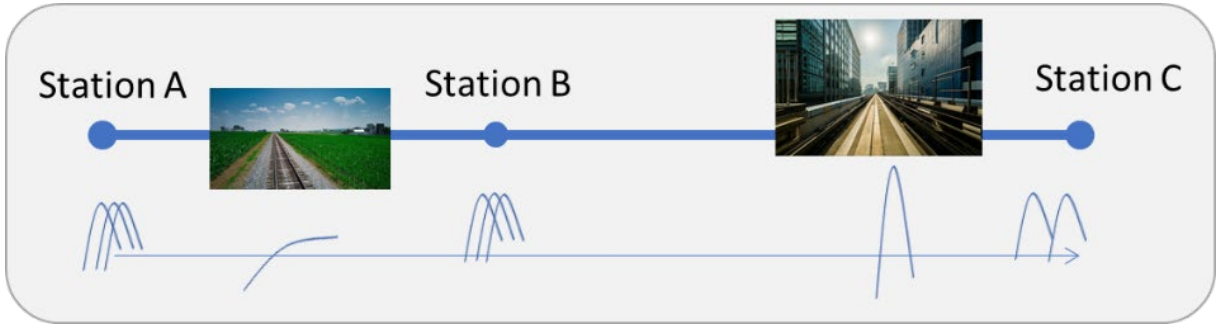
OBJECTIVE

To provide an end-to-end chain capable of simulating and evaluating realistic GNSS reception conditions **function of time and all along a railway line**

A TWO-STEP METHODOLOGY

1. Data-driven characterization of the reception environment

2. Error modelling for each type of environment



WHAT IS THE AVAILABLE INFORMATION ?

Using GNSS Raw measurements (RINEX)

Data

Small
- C/N0
- Elevations
- Nb satellites
- PDOP-GDOP

Medium
- …

+
- Low order statistical moments
- Delays Iono
- Delays Tropo

Large
- …

+
- High order statistical moments
- Multi-frequencies
- Multi-constellations

Dimension of the problem ↗

Classes

Primary classes
- *Buildings*
- *Tree*
- *Open-sky*
- *Bridge*

Secondary classes
- *train-station,*
- *mixed_tree_open*
- *mixed_tree_build*
- *mixed_build_open*

WHAT IS THE AVAILABLE INFORMATION ?

Using public sources

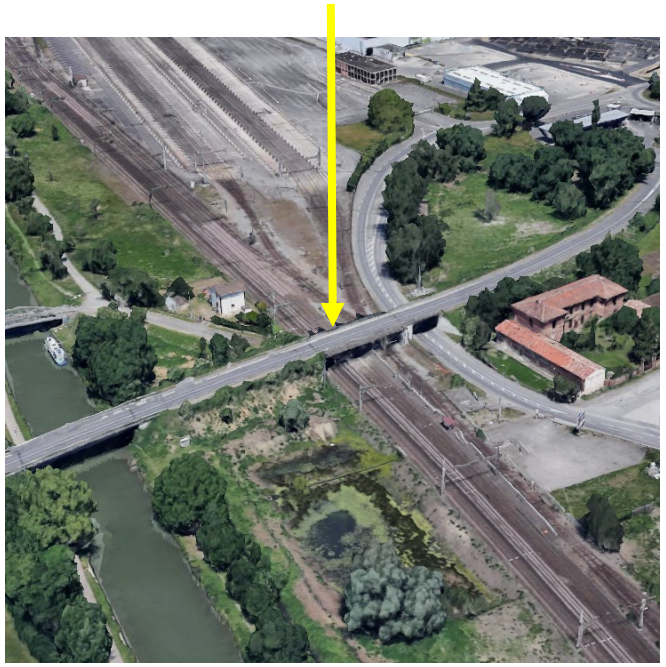
Buildings

Infrared

Tree

Satellite

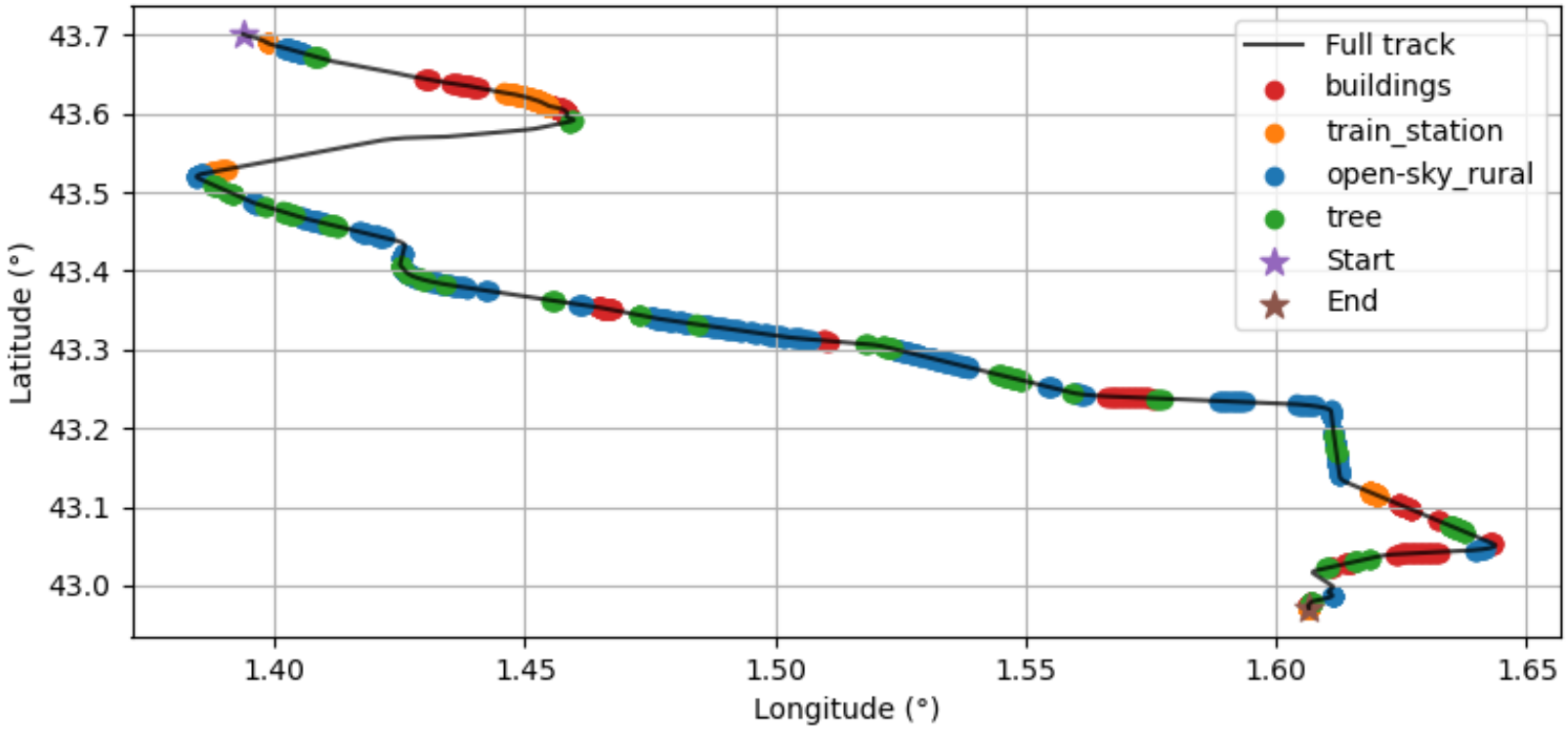
Tunnels



Google Earth / Copernicus

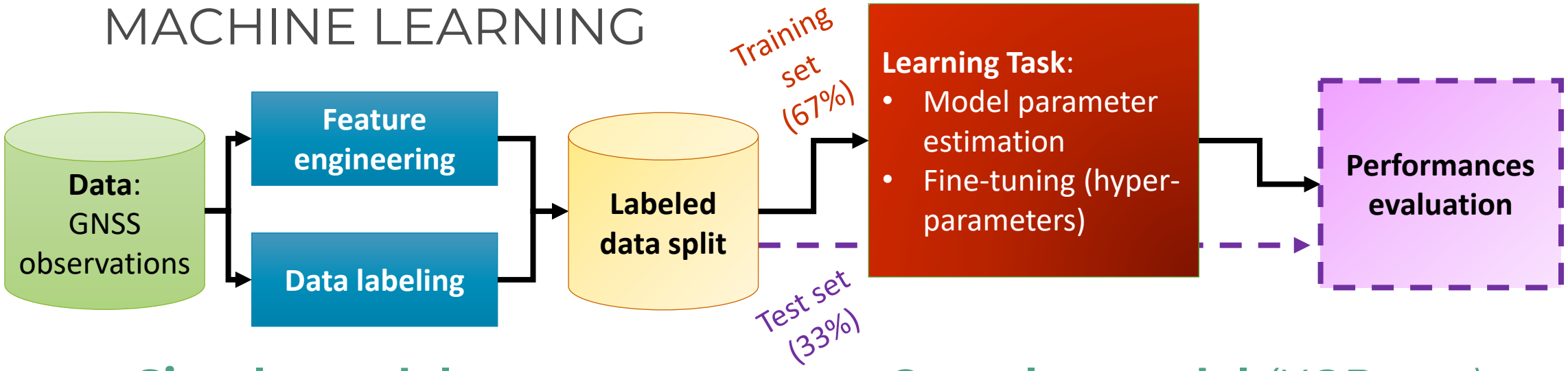
WHAT IS THE AVAILABLE INFORMATION ?

After the labelling process (CLUG dataset)



Majority of mixed classes

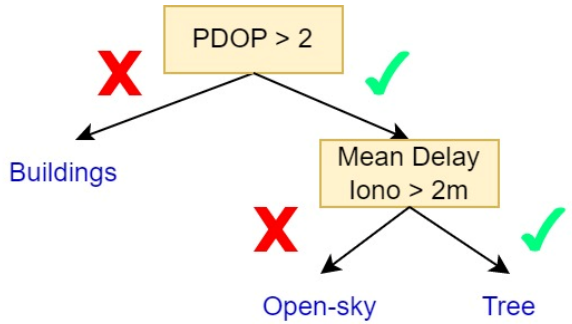
MACHINE LEARNING



- **Simple model** *(Multiclass Logistic Regression)*:
  - Linear model
  - Easy to interpret
  - Lower performance

$$p_i = \frac{1}{1 + e^{\beta x_i + \beta_0}}$$

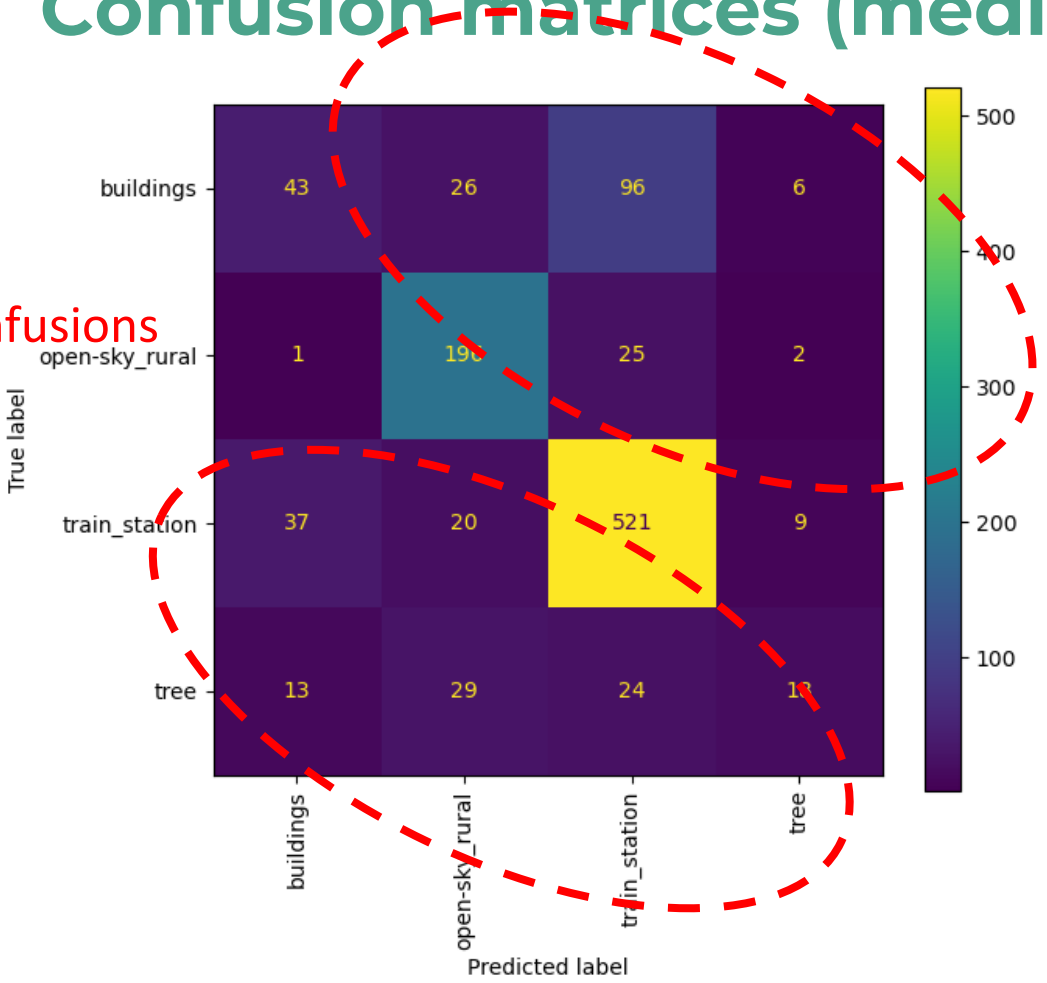
- **Complex model** *(XGBoost)*
  - Boosting methods based on tree classifier
  - Hard to interpret
  - No assumption of linearity



MACHINE LEARNING

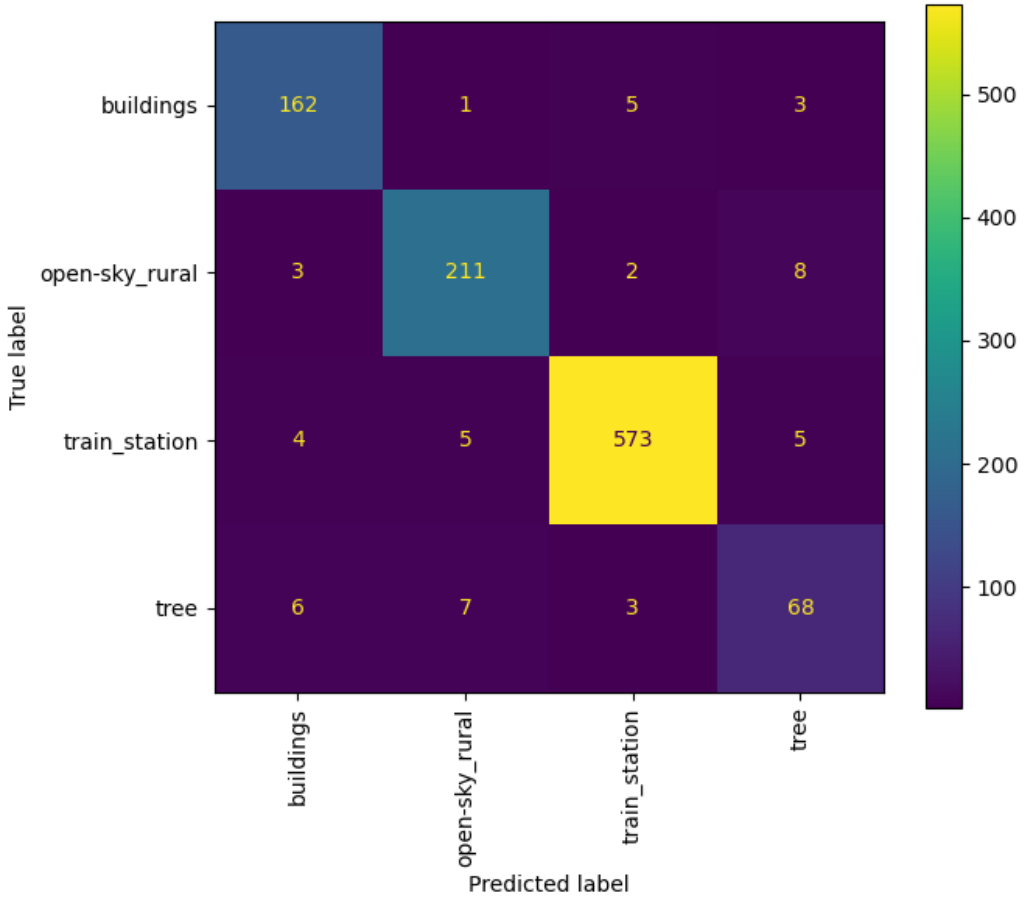
Confusion matrices (medium dataset ~ low dim)

Confusions



Linear Model

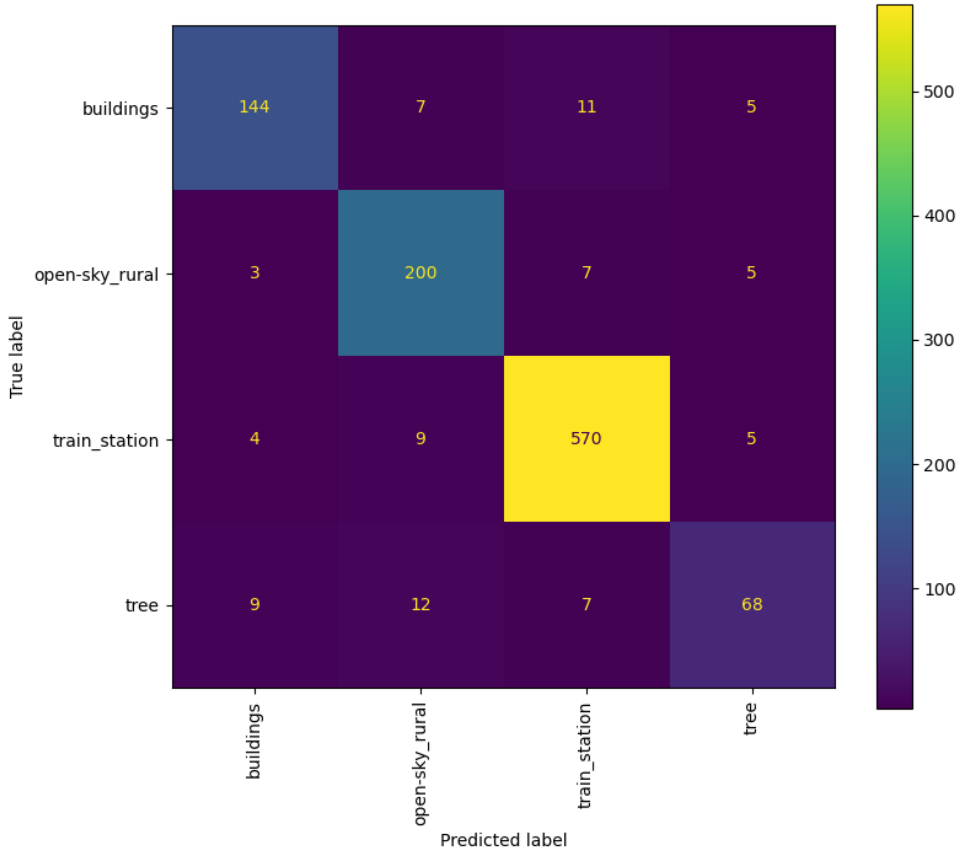
Accuracy = 0.73



Non-Linear Model

Accuracy = 0.95

Confusion matrices (large dataset ~ high dim)



≈

Linear Model

Accuracy = 0.92

But <u>no</u> confidence about predictions at future times

Non-Linear Model

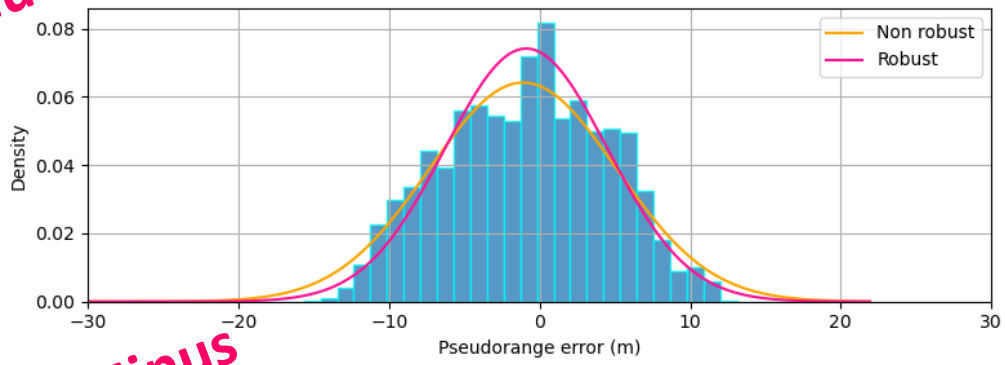
Accuracy = 0.99

ERROR MODELING

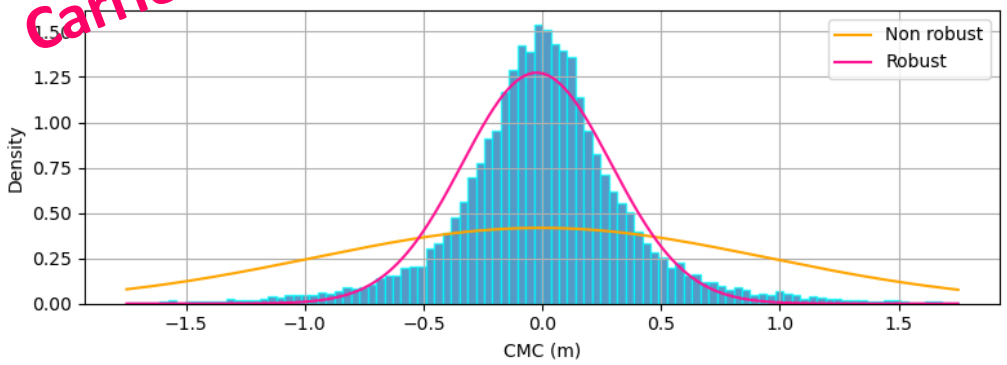
Some insights (GPS L1)

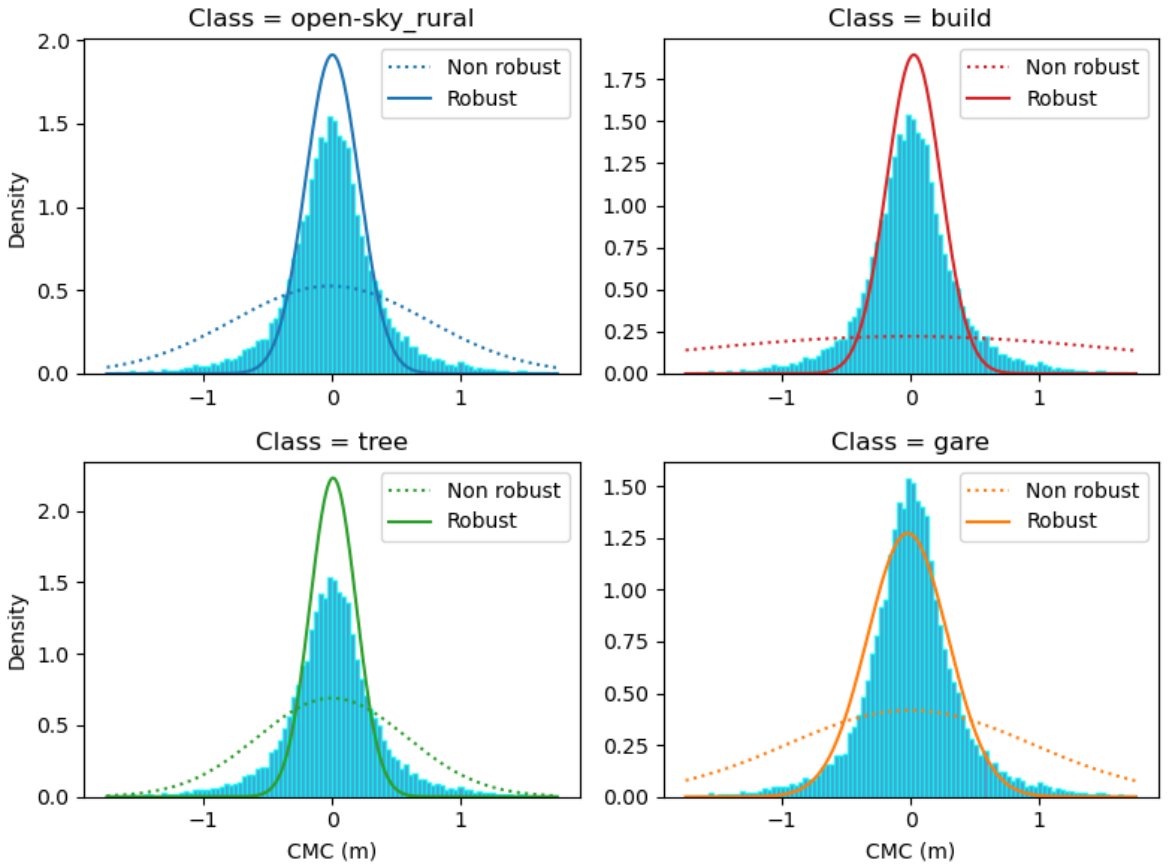
Need of Robust Gaussian approaches
(ex: *Minimum Covariance Determinant*)

Pseudoranges



Code Minus Carrier





SOME FEEDBACK ON THE PROBLEM

On the environment choice

- Multiple choices depending on the source of information
- Little work done on the temporal variability

On the machine learning aspect

- Strong correlations between observations (environments = “groups”)
- How to prevent the model to learn the spatial information ???

YOUR CONTACTS

Juliette MARAIS
Uni. Gustave Eiffel

Juliette.marais@univ-eiffel.fr

Quentin MAYOLLE
Railenium

Quentin.mayolle@railenium.eu

RAILENIUM
RAIL RESEARCH & INNOVATION

RAILENIUM
180 rue Joseph-Louis Lagrange
59308 VALENCIENNES Cedex

www.railenium.eu

i-TRANS