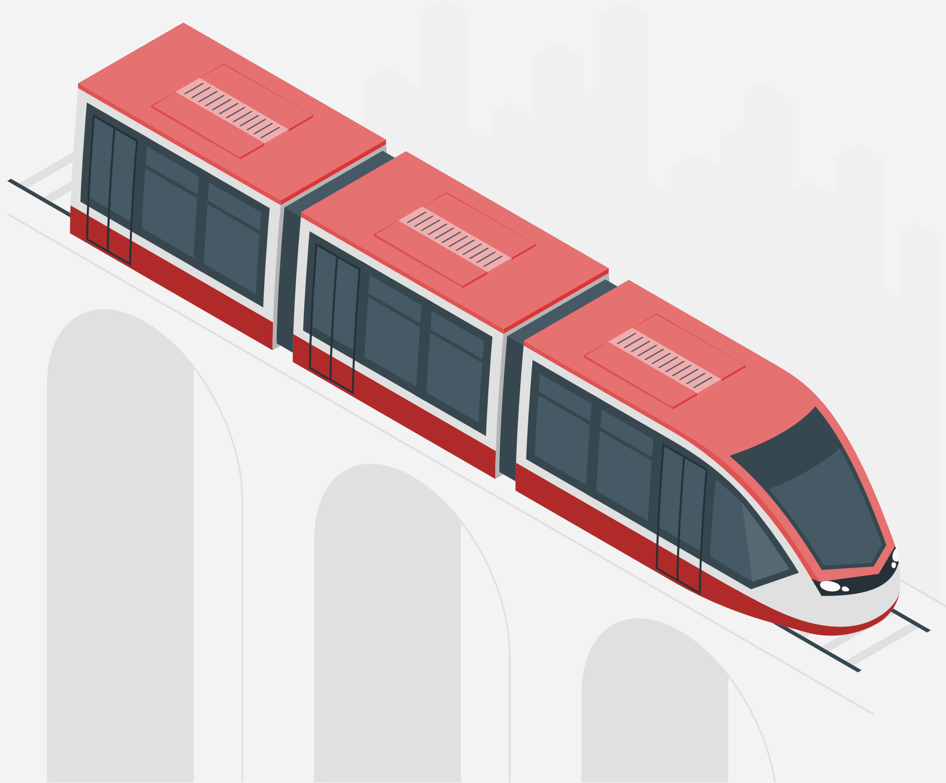
A Deep Learning-based Anomaly Detection System for Railway Equipment

Michele Bevilacqua (RFI), Walter Didimo (UniPG), Alessandro Eugeni (UniPG), Fabrizio Montecchiani (UniPG), Giovanni Spinelli (RFI)

RFI
RETE FERROVIARIA ITALIANA
GRUPPO FERROVIE DELLO STATO ITALIANE

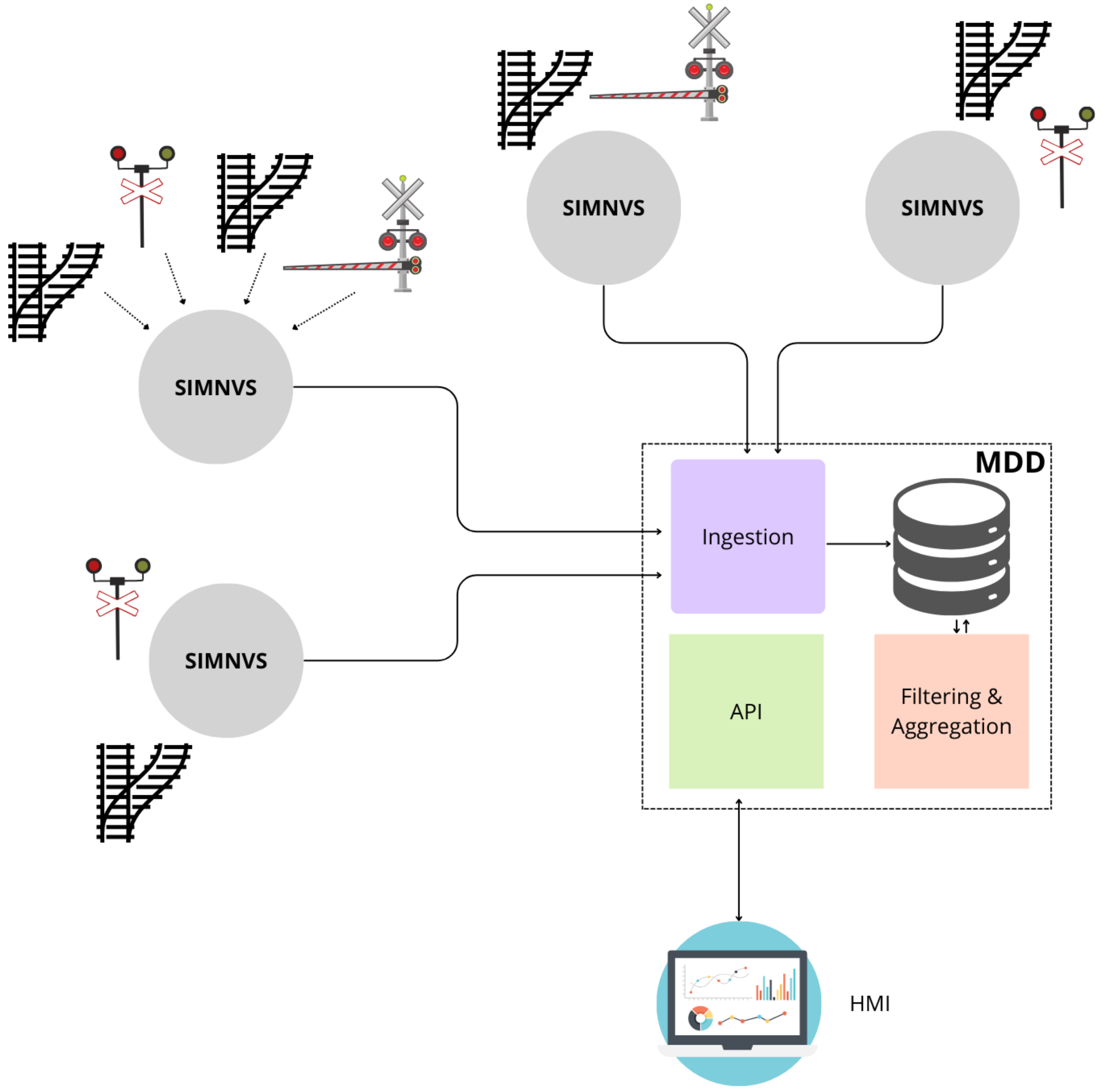
A.D. 1308
unipg
DIPARTIMENTO
DI INGEGNERIA



ITA DATA 2025

Context

- Established collaboration between RFI and CINI.
  - Our research is embedded in a larger project aimed at developing a software platform to ingest, store, and analyze diagnostic data from the railway network.



Motivation

- In 2024, for high-speed trains:
  - 57% of delays caused by infrastructural problems
  - 3.16 billion euros lost due to delays for companies and travelers

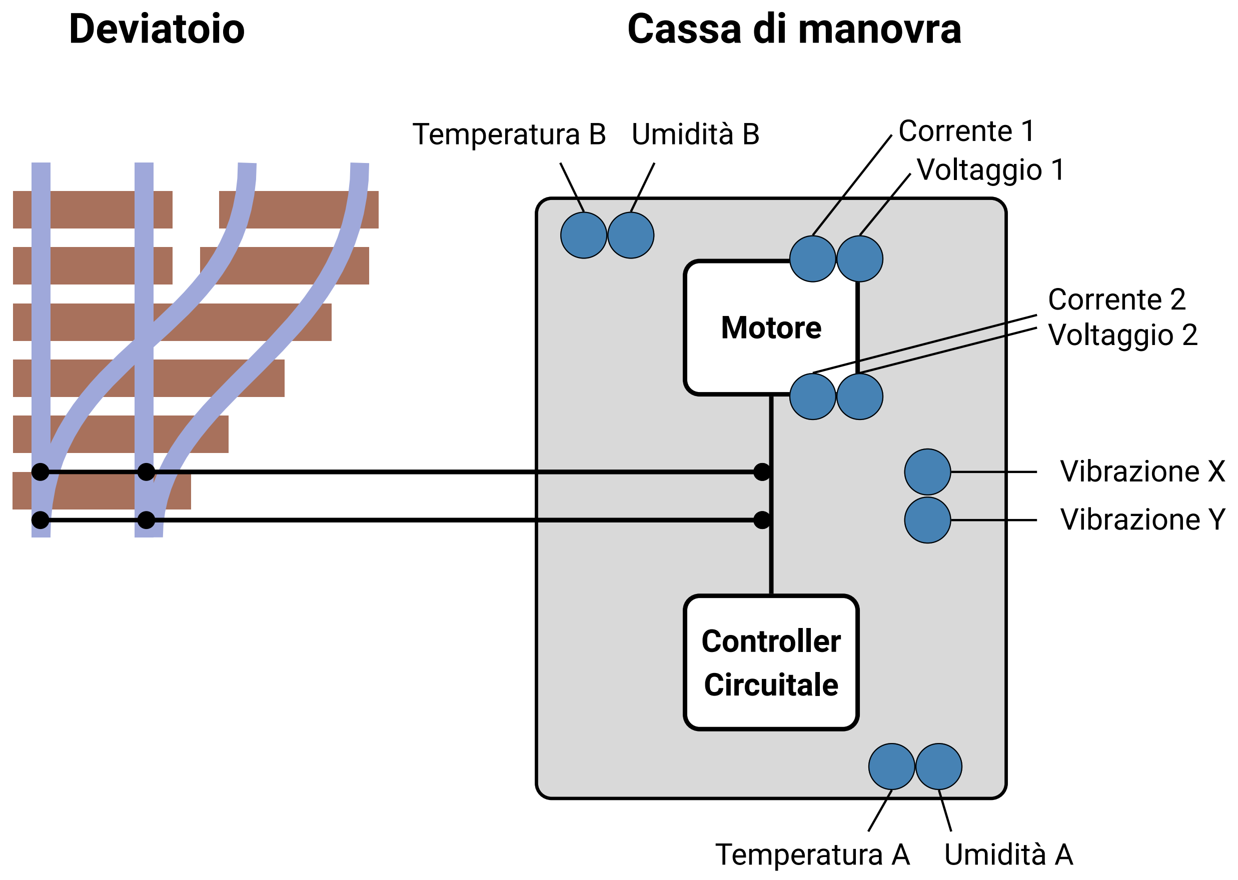
Sources: RFI, Unimpresa

Contribution

- We focus on a specific type of device, namely railway switches
  - critical assets: faults can freeze train lines

- Objective: develop a data-driven solution for predictive maintenance of railway switches

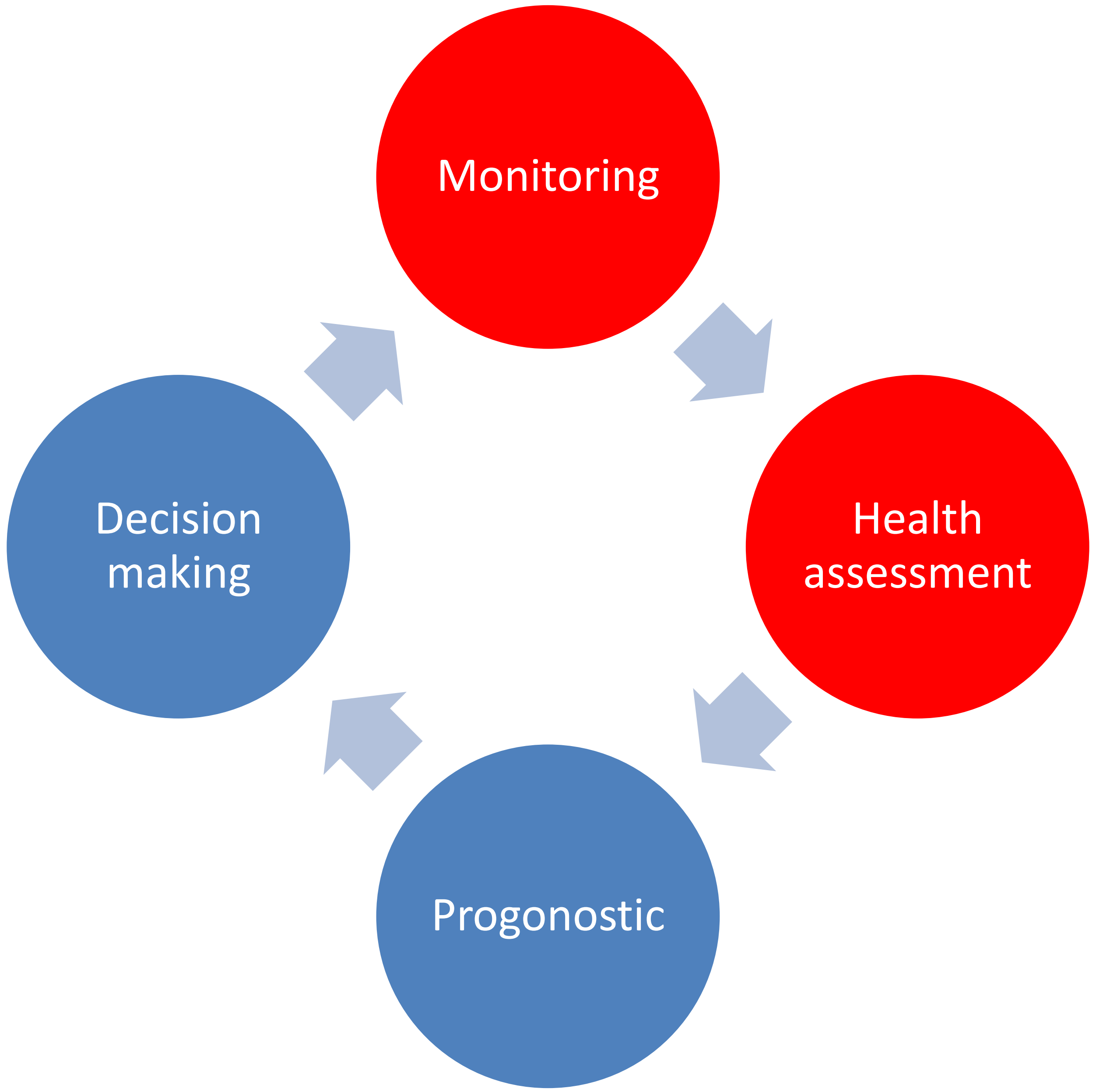
- Twofold contribution:
  - Original methodology for the creation of a hybrid dataset (real and synthetic data), with various degradation levels.
  - Thorough experimental evaluation of SOTA deep-learning models to extract a health indicator index of a railway switch.



Condition-Based Maintenance (CBM)

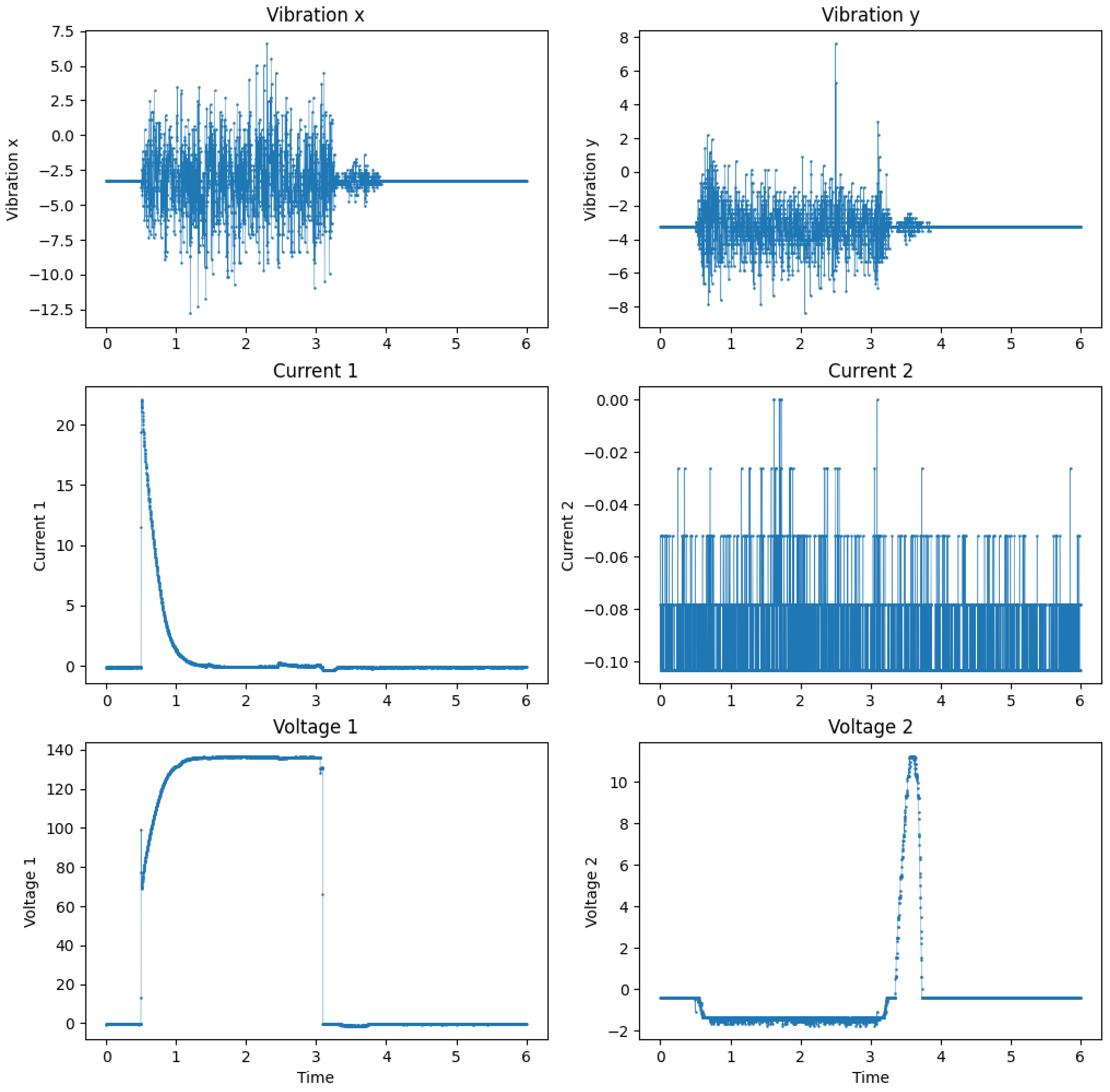
- Maintenance based on asset health, rather than periodic schedules
  - Focus on health assessment

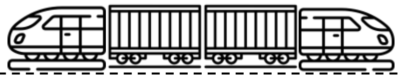
- Data-driven approaches avoid explicit modeling of physical assets



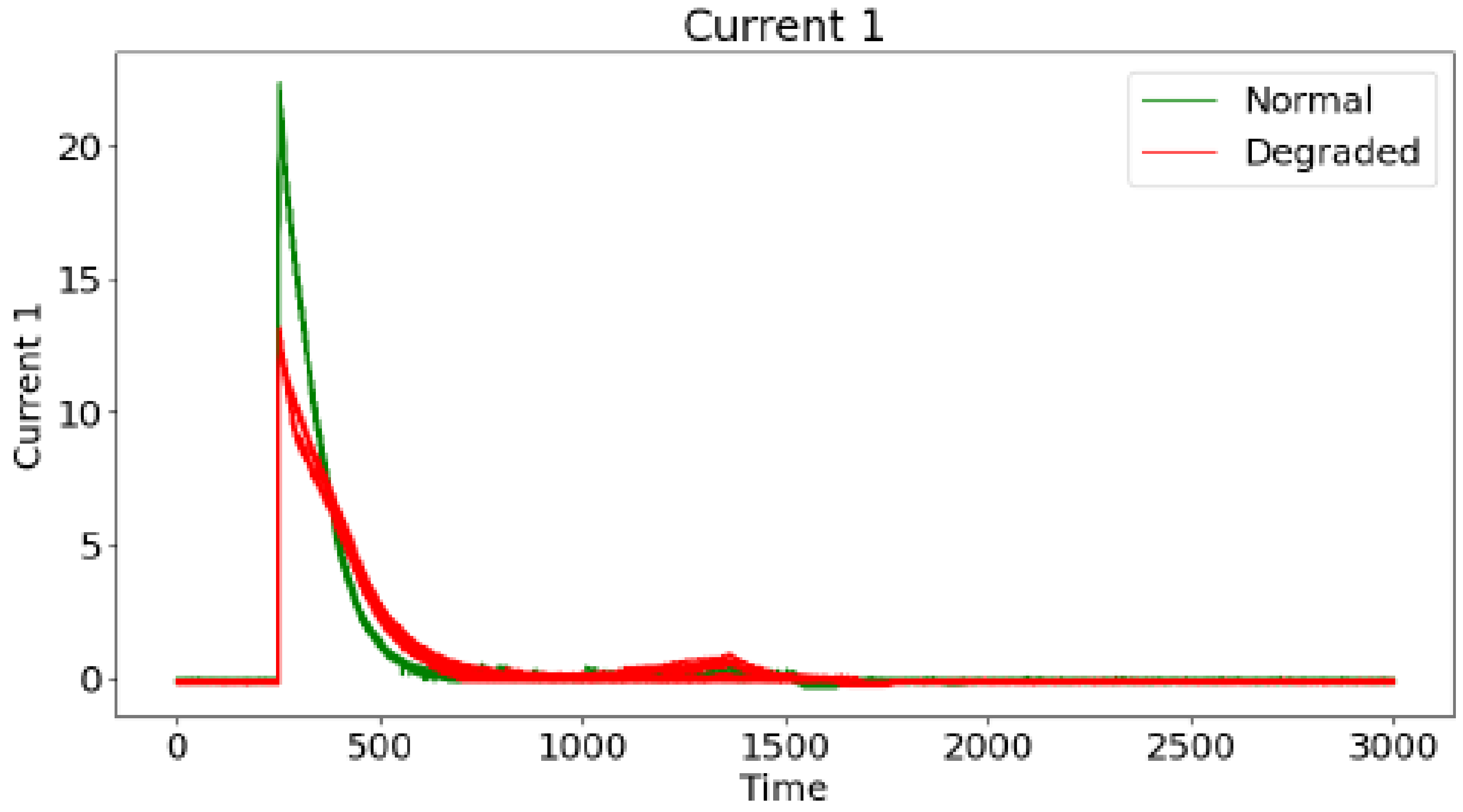
Collecting the Real dataset

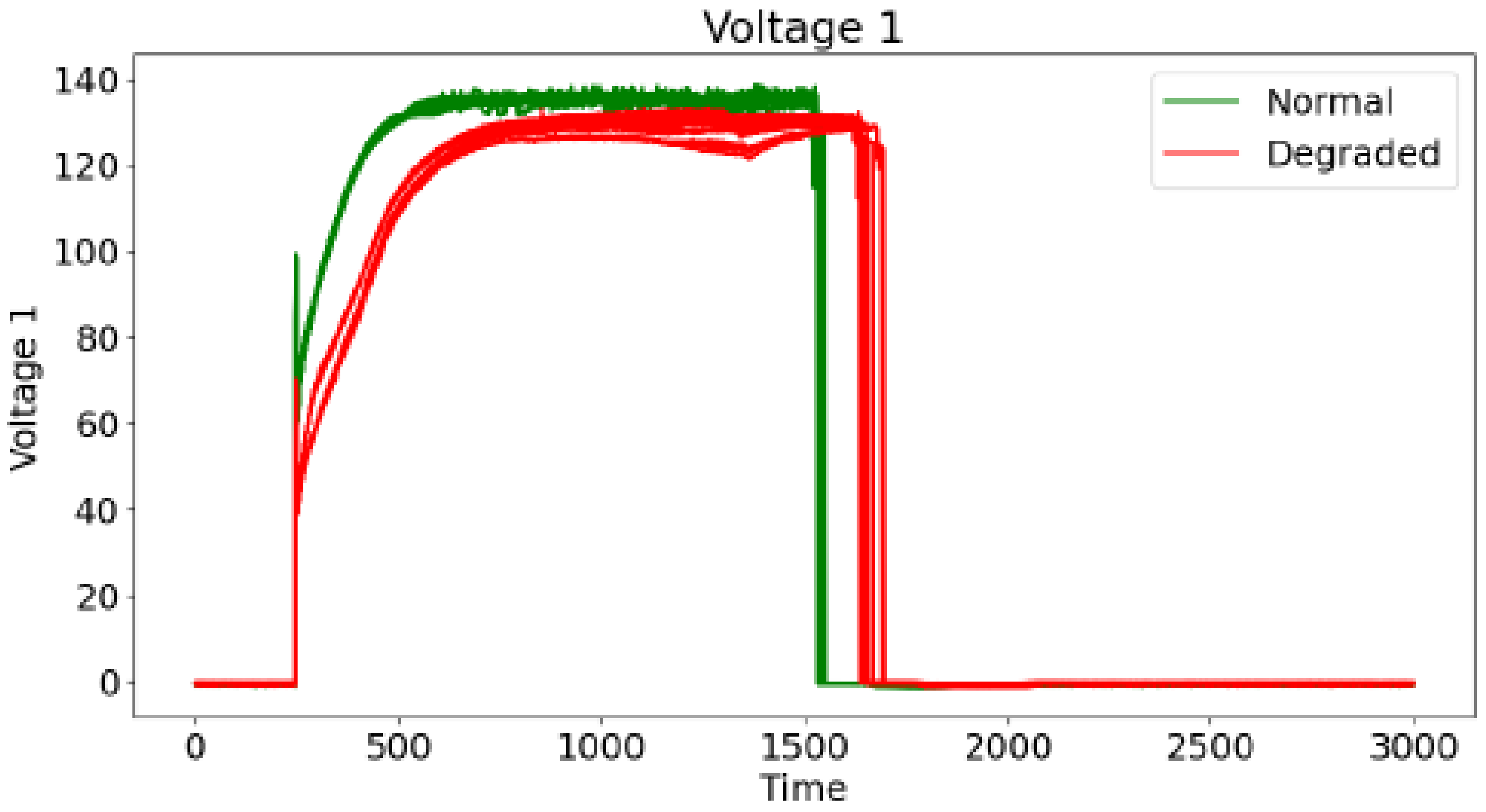
- Sensors: current, voltage, vibration, temperature, humidity
- Real dataset:
  - 153 maneuvers (time series)
    - 116 normal
    - 37 abnormal
  - Anomalies injected via mechanical resistance (friction)
    - Simulates aging or absence of lubricant of the physical asset.





Collecting the Real dataset

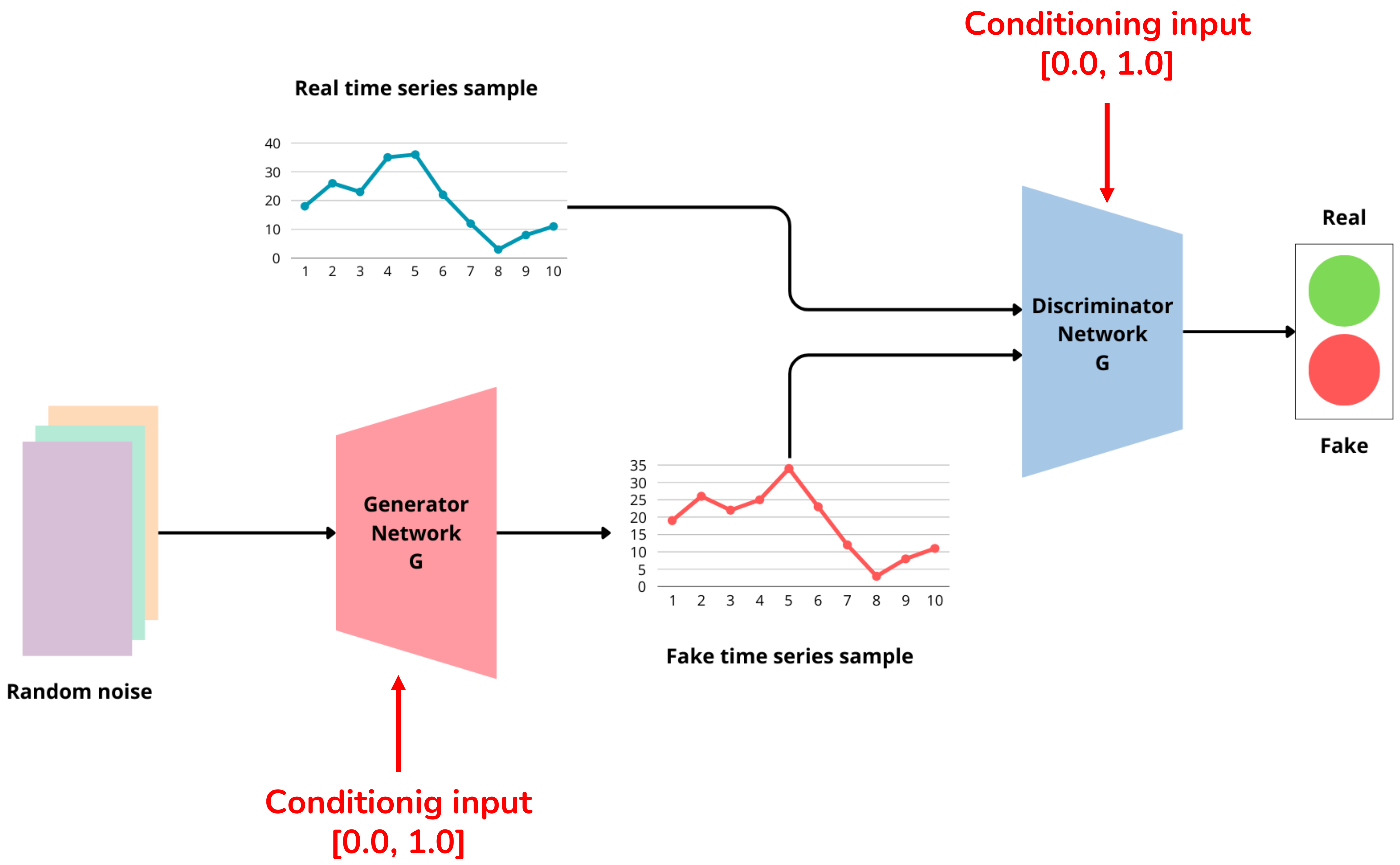




Computing synthetic data

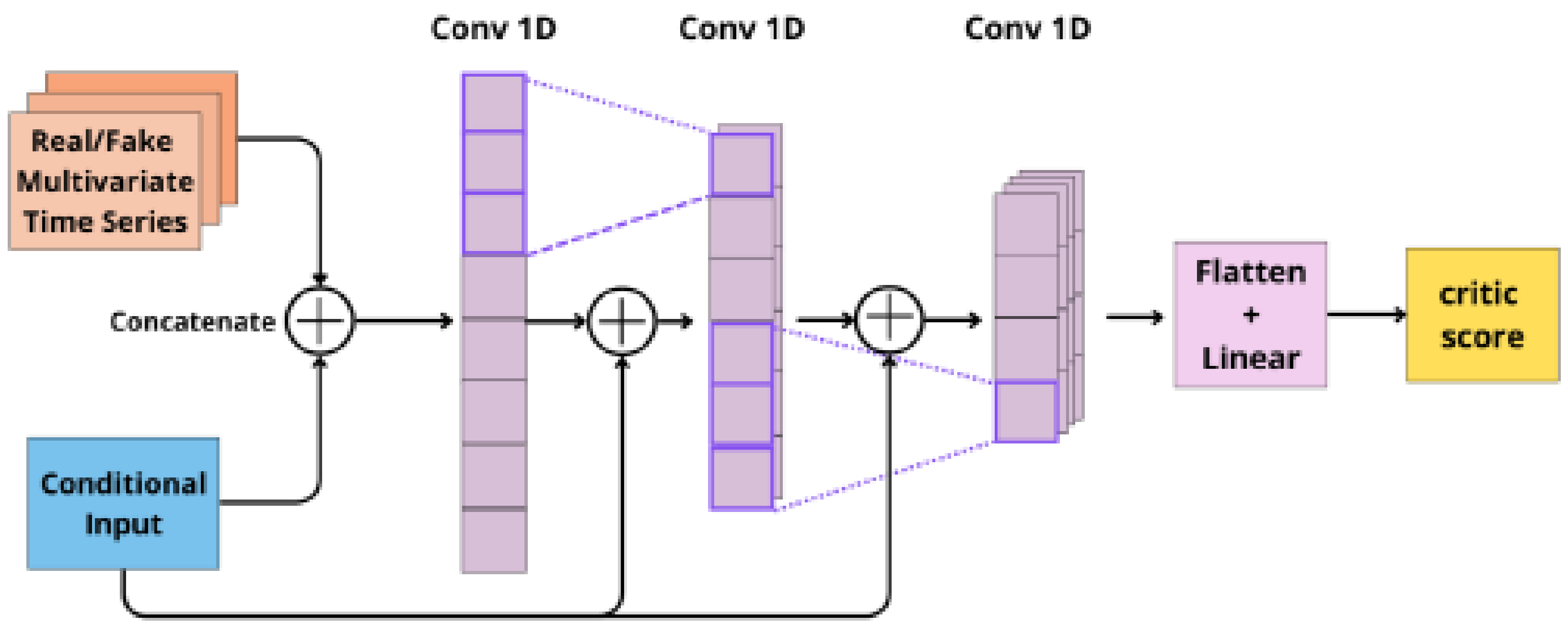
- Issues of real dataset:
  - small,
  - only extremal behaviors

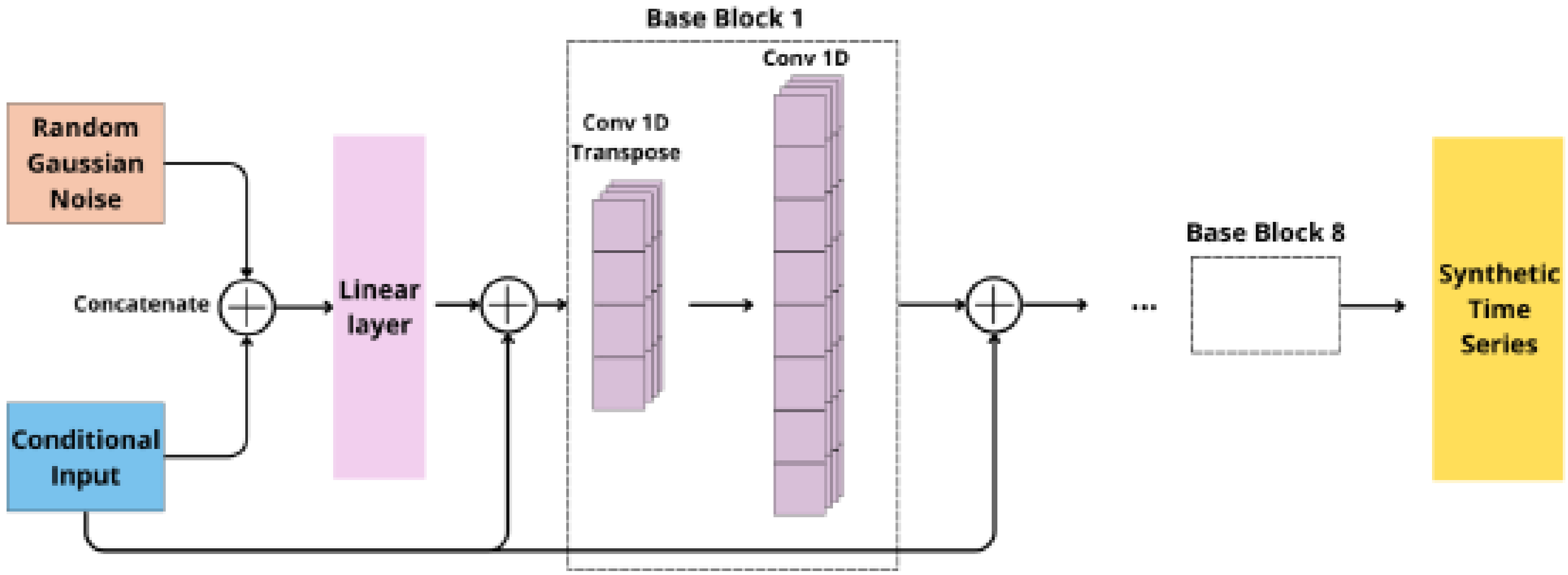
- Solution:
  ad-hoc GAN architecture to deal with multiple degradation levels
  (0.1 → 0.9)



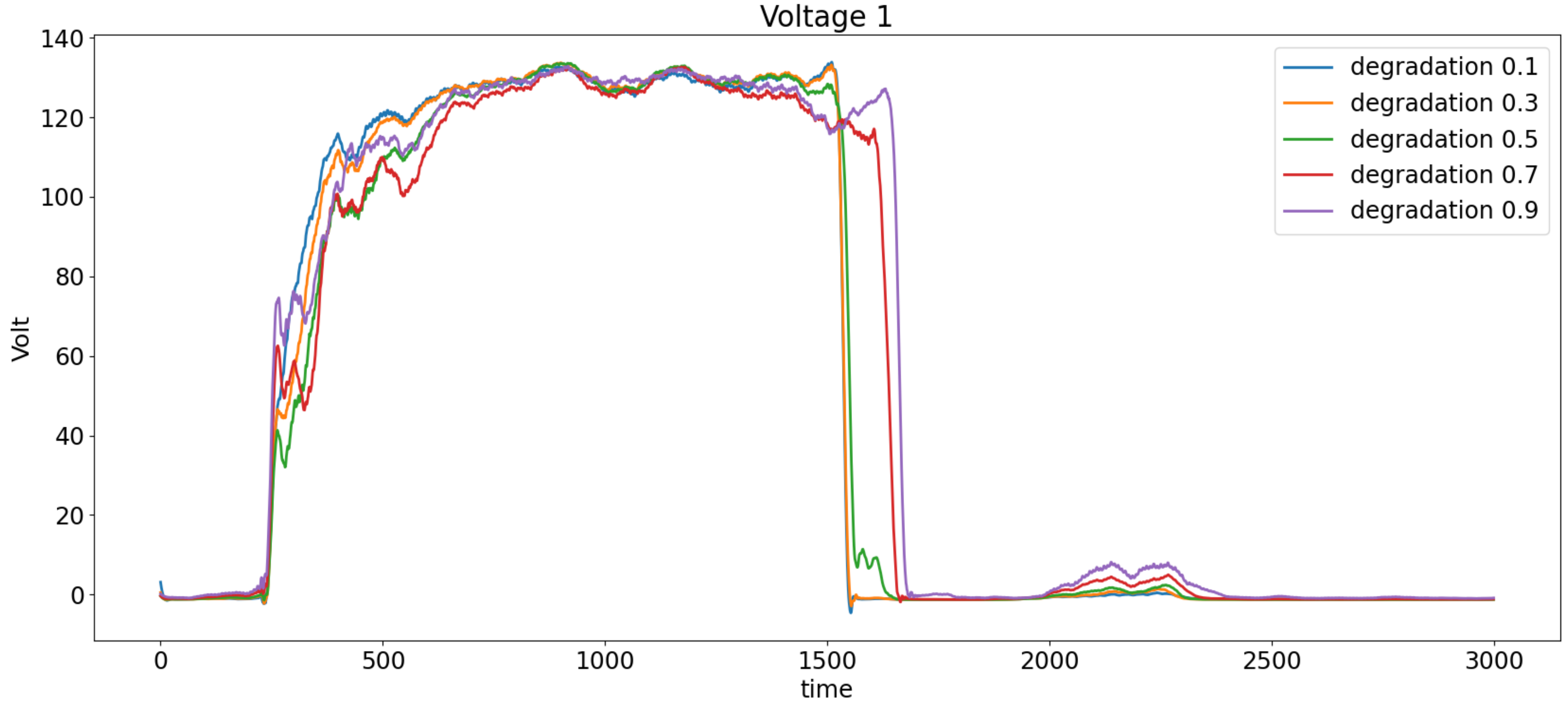
Conditional Wasserstein GAN with Gradient Penalty

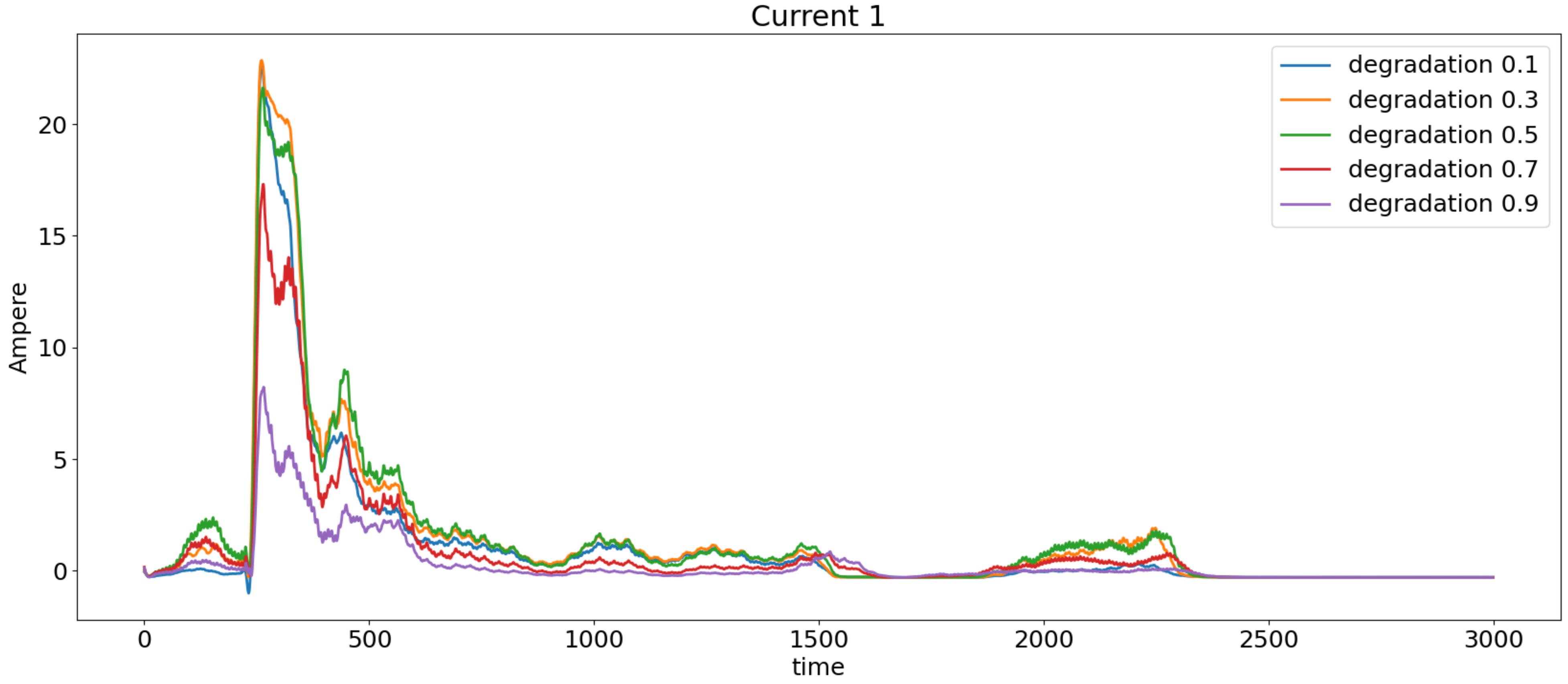
- More robust training with fewer chances of mode collapse
- Conditioning input to interpolate between extremal behaviors





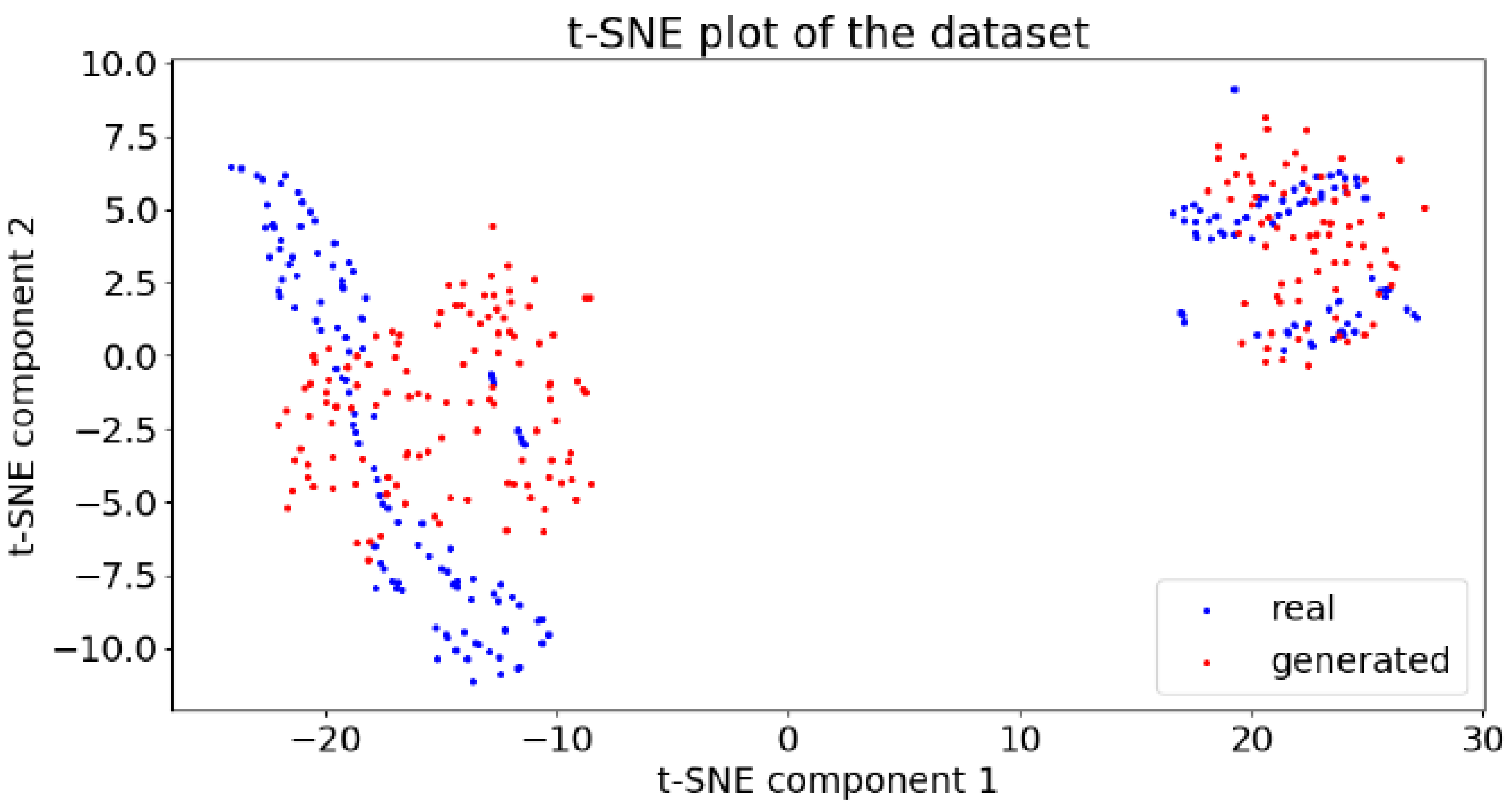
Synthetic dataset (1,250 maneuvers)





Validation

- t-SNE plot
  (reconstruction of the real dataset)
- Train-Synthetic Test-Real (omitted)
- Discussions with RFI engineers

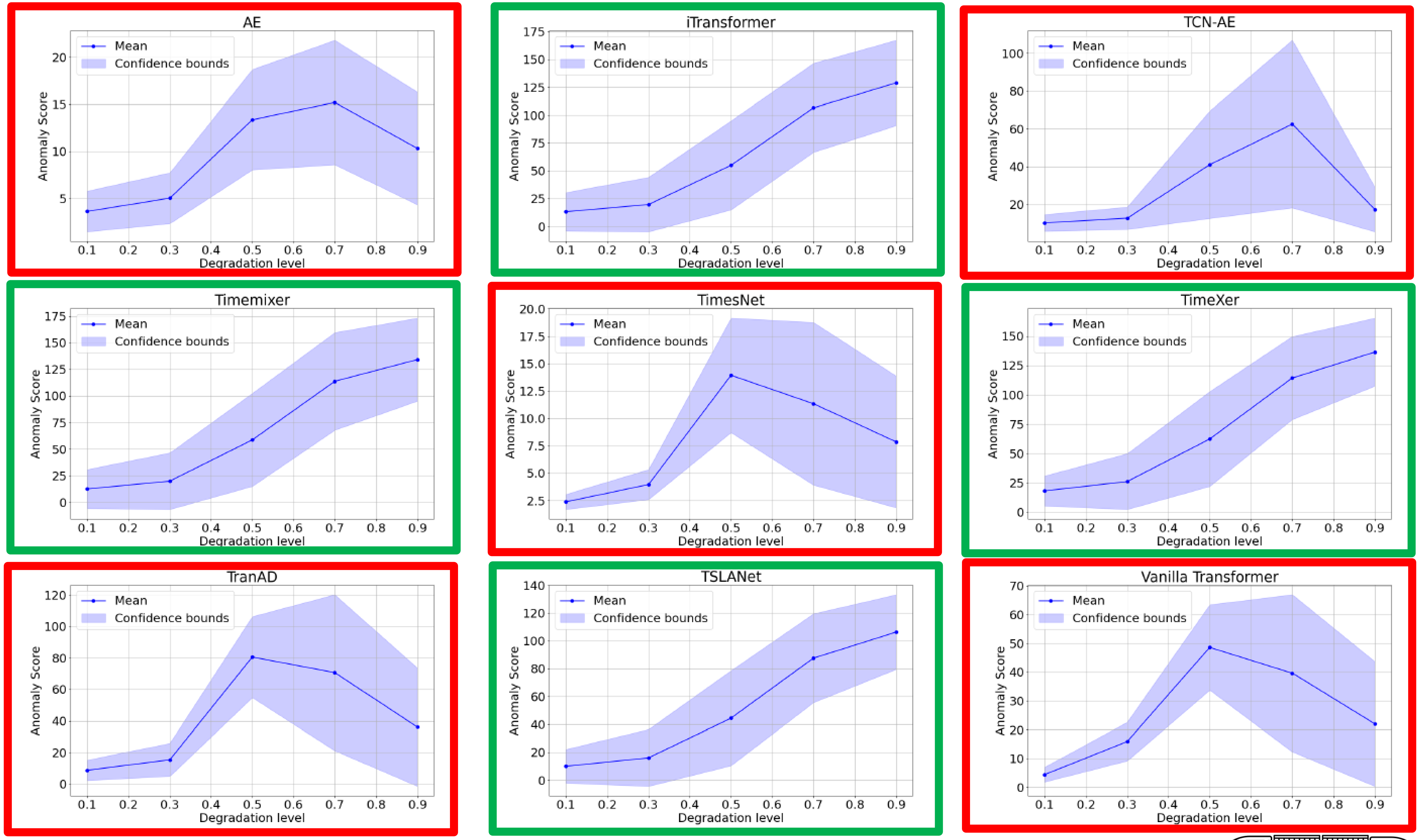


Health indicator index

- Approach: we use the mean square error (MSE) of SOTA reconstruction-based DL models
  - Standard approach when dealing with time-series

- Candidate models identified through a preliminary analysis:
  - iTransformer (transformer encoder-only)
  - TimeMixer (fully MLP with scale decomposition)
  - TimeXer (transformer encoder-only)
  - TSLANet (convolutional neural network)
  - TranAD (transformer + GAN)
  - TimesNet (based on ResNet)
  - Baselines:
    - AutoEncoder
    - Vanilla Transformer

Results



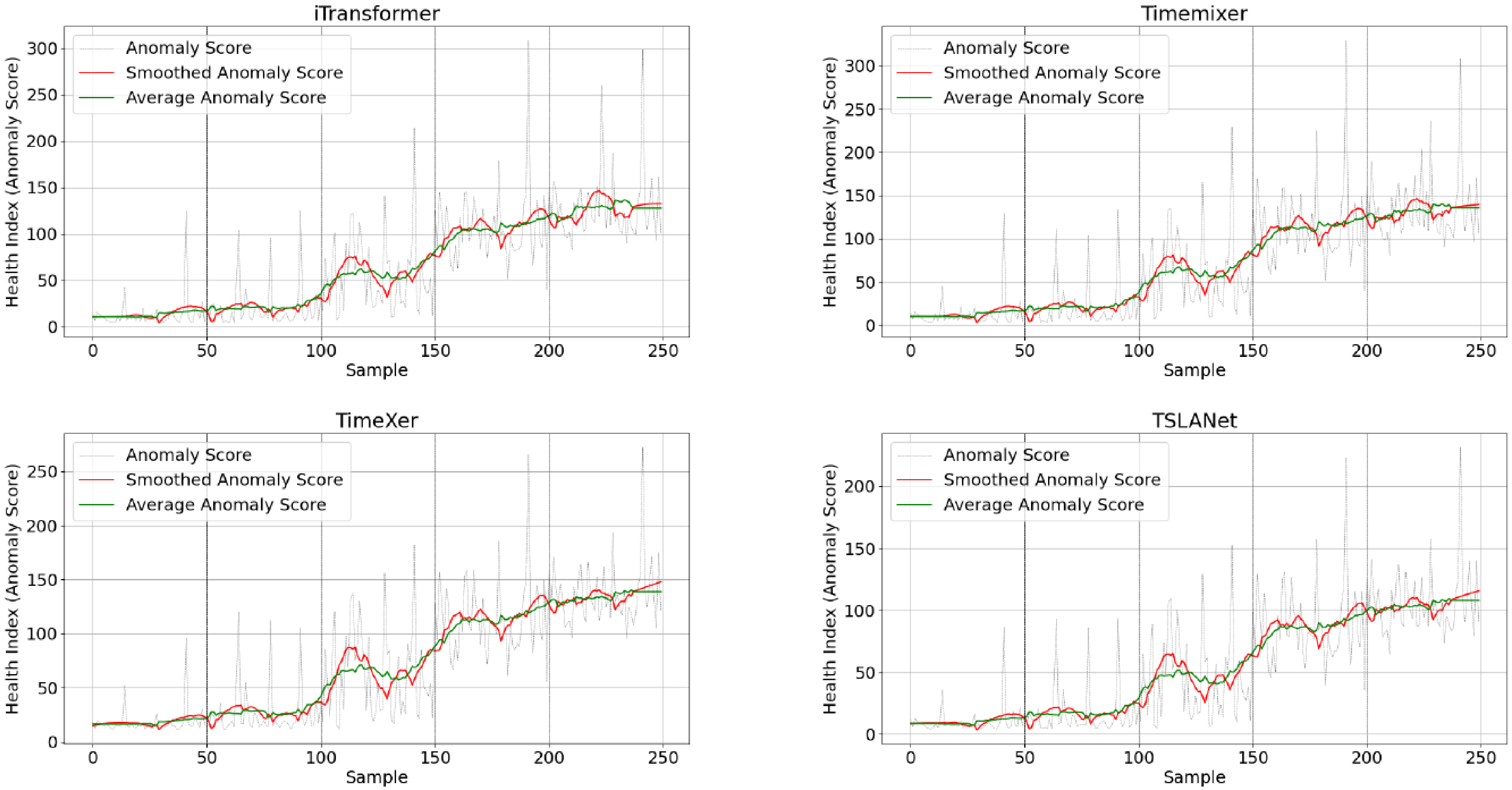
Training performed on common laptop --Ryzen 7 6800HS CPU, 32 GB of RAM, Nvidia GeForce 3050

Results

- Robust & monotonic trends achieved by:
  - iTransformer, TimeMixer, TimeXer, TSLANet

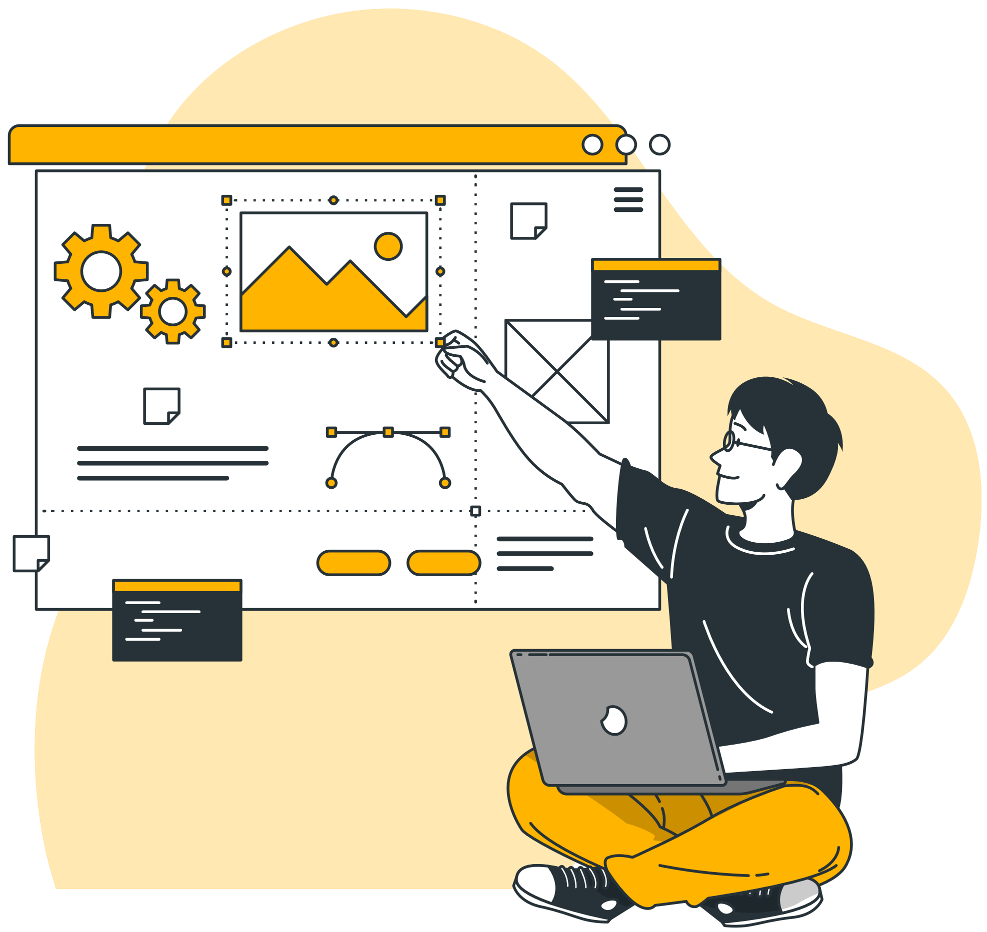
- All other models achieved inconsistent results

Health index evolution over the samples:
The degradation level is 0.1 for the first maneuver and is incremented by 0.2 every 50 samples.



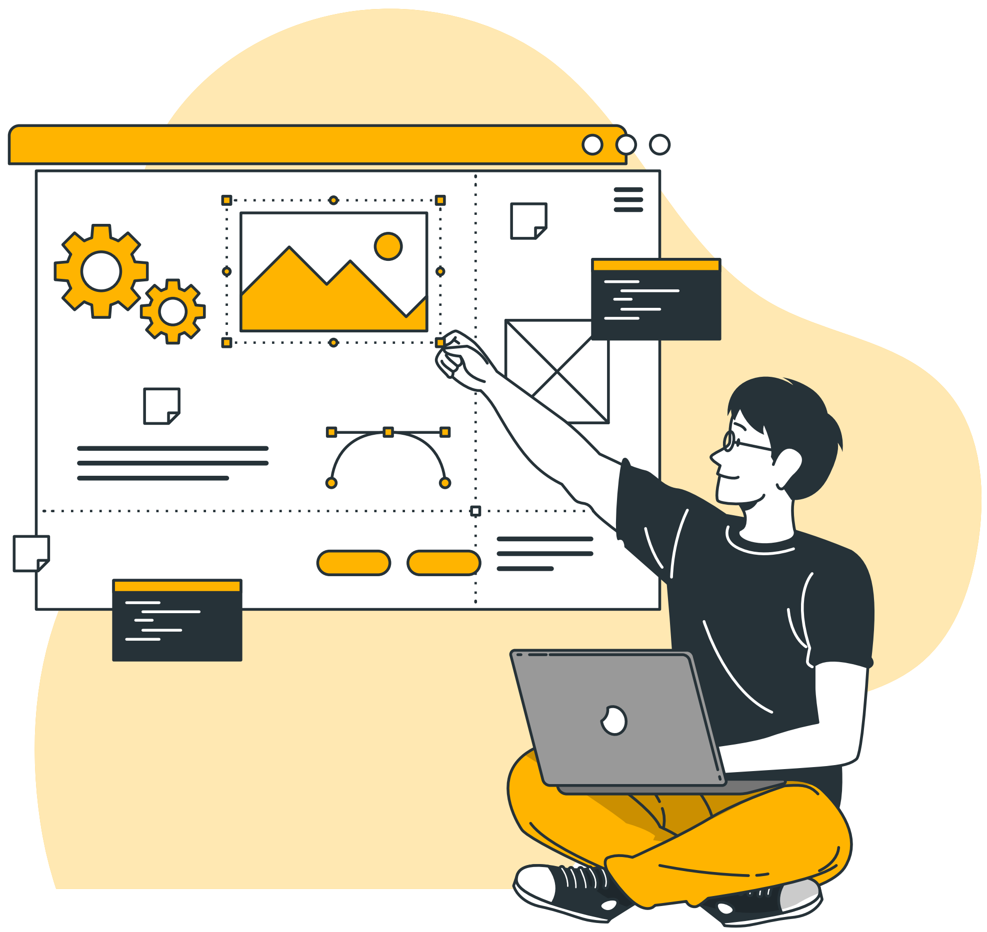
Future work

- More general dataset and include exogenous variables
- Run-to-failure data collection,
- Estimation of remaining useful life (RUL),
- Work on explainability of models



Future work

- More general dataset and include exogenous variables
- Run-to-failure data collection,
- Estimation of remaining useful life (RUL),
- Work on explainability of models



THANK YOU FOR YOUR ATTENTION!