Generation of Synthetic SCADA Signals using cGANs for Enhanced Wind Turbine Fault Detection and Prognosis

Ali Eftekhari Milani

Donatella Zappalá

Francesco Castellani

Simon J. Watson

7th Wind Energy Systems Engineering Workshop

TU Delft

2024-12-04

Outline

- Context
- Fault detection
- Problem
- Proposed solution
- Methodology
- Results
- Conclusions and future work

TU Delft

Context of the study

O&M costs are significant!
Esp. offshore (up to 30% of total cost)

Reactive Maintenance
- Fix when it breaks

Preventive Maintenance
- Maintain at regular intervals so it does not break

Predictive Maintenance
- Predict when it will break and maintain accordingly

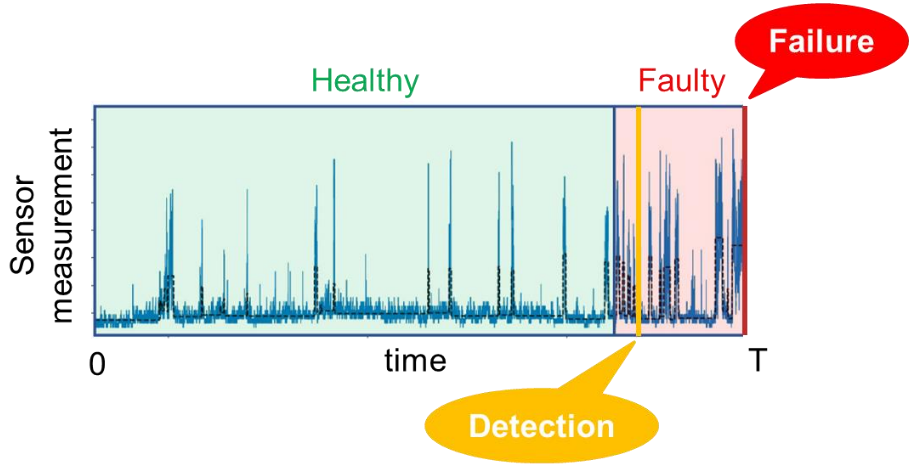
Faults need to be predicted well before a failure happens

TU Delft

Fault detection: Two main approaches

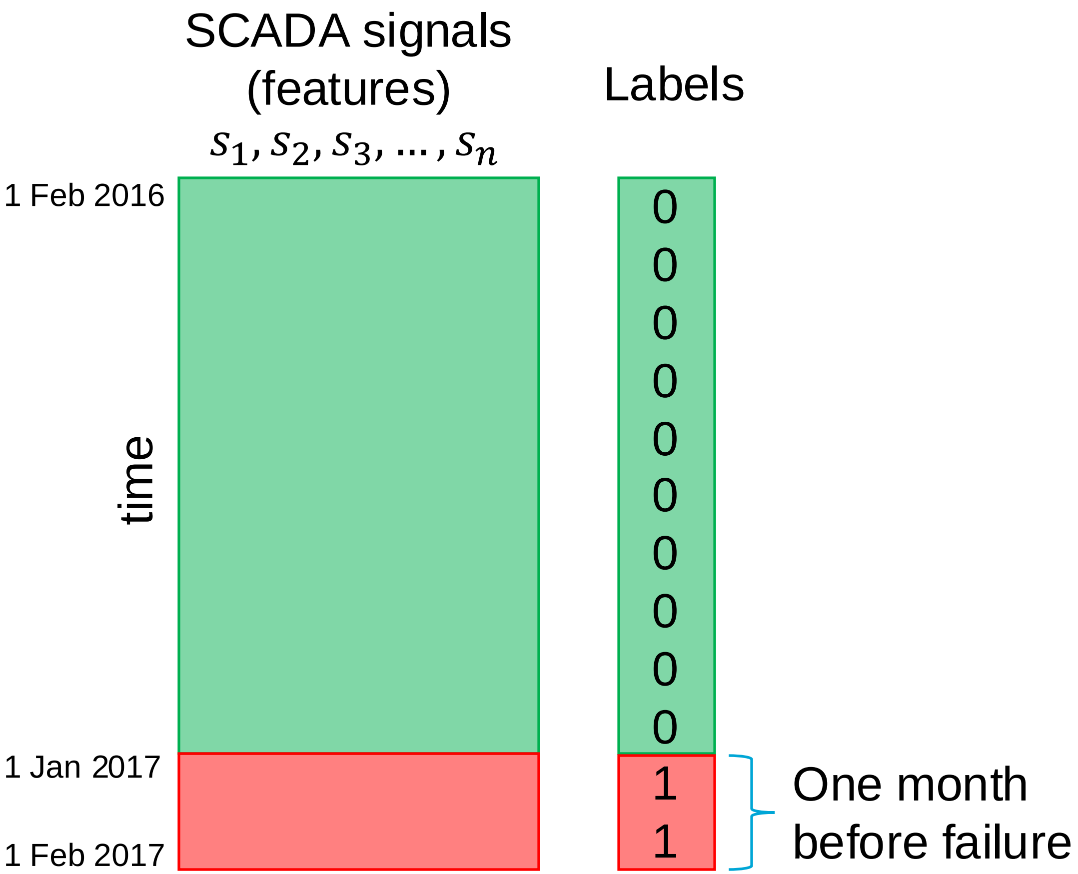
- Based on Normal Behaviour Modelling (NBM)
  - \+ Only healthy data needed for training
  - – Cannot distinguish between different failure types

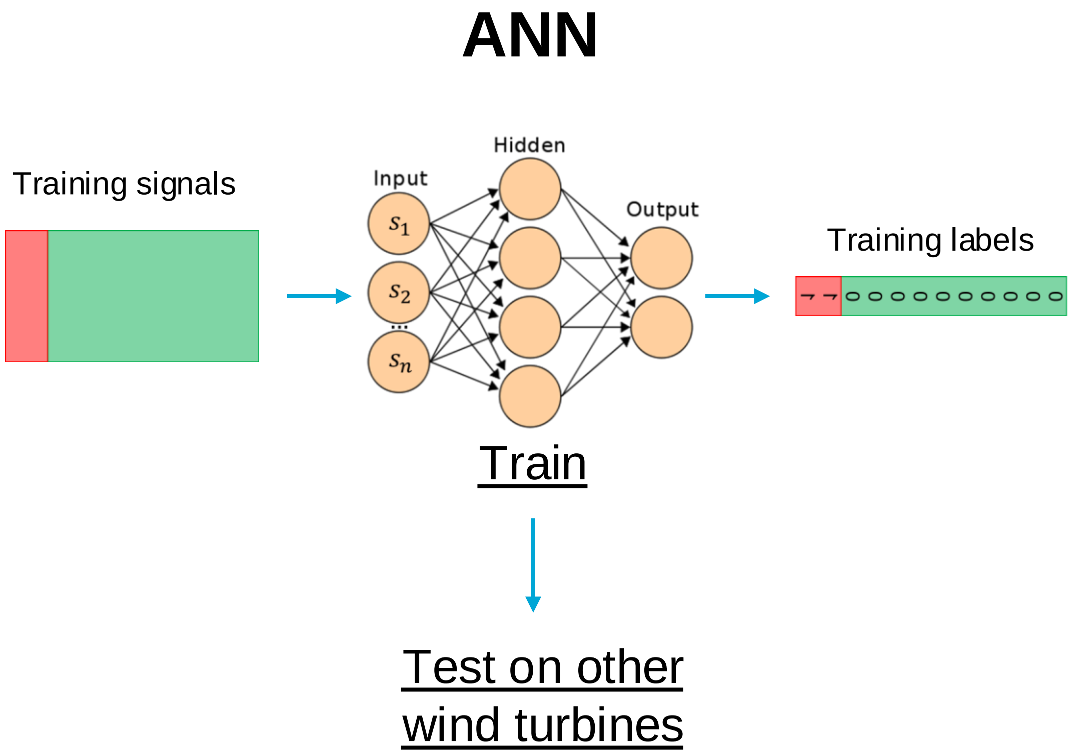
- Based on Classification
  - \+ Different failure types can be distinguished
  - – Both healthy and faulty data needed for training



When detection of <u>a certain fault type</u>, e.g., Gearbox fault, is required → Classification ~~Normal Behaviour Modelling~~

Classification





Problem

Failures are rare events!

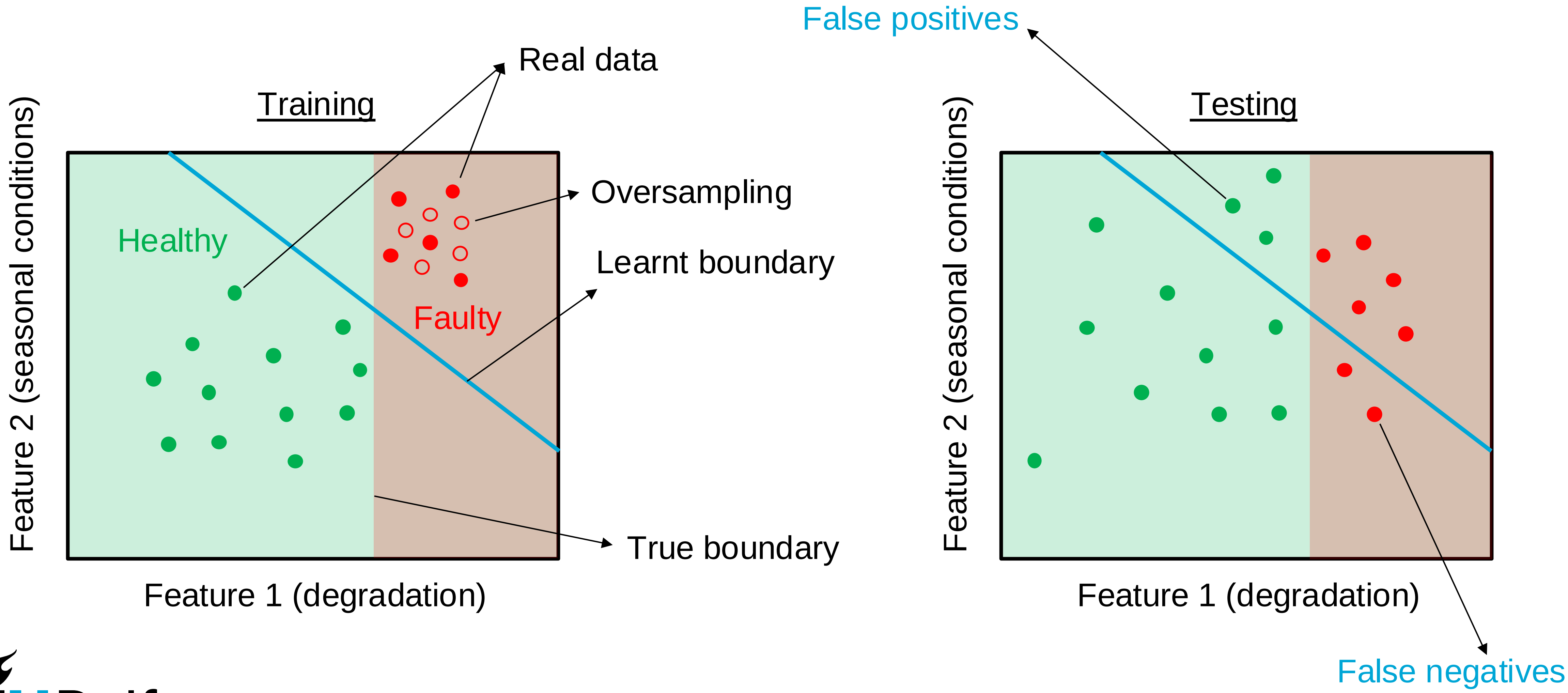
Available datasets:

Usually a handful of turbines
during a few years of operation

Very few failures events in the data

→ This leads to problems when using classification methods

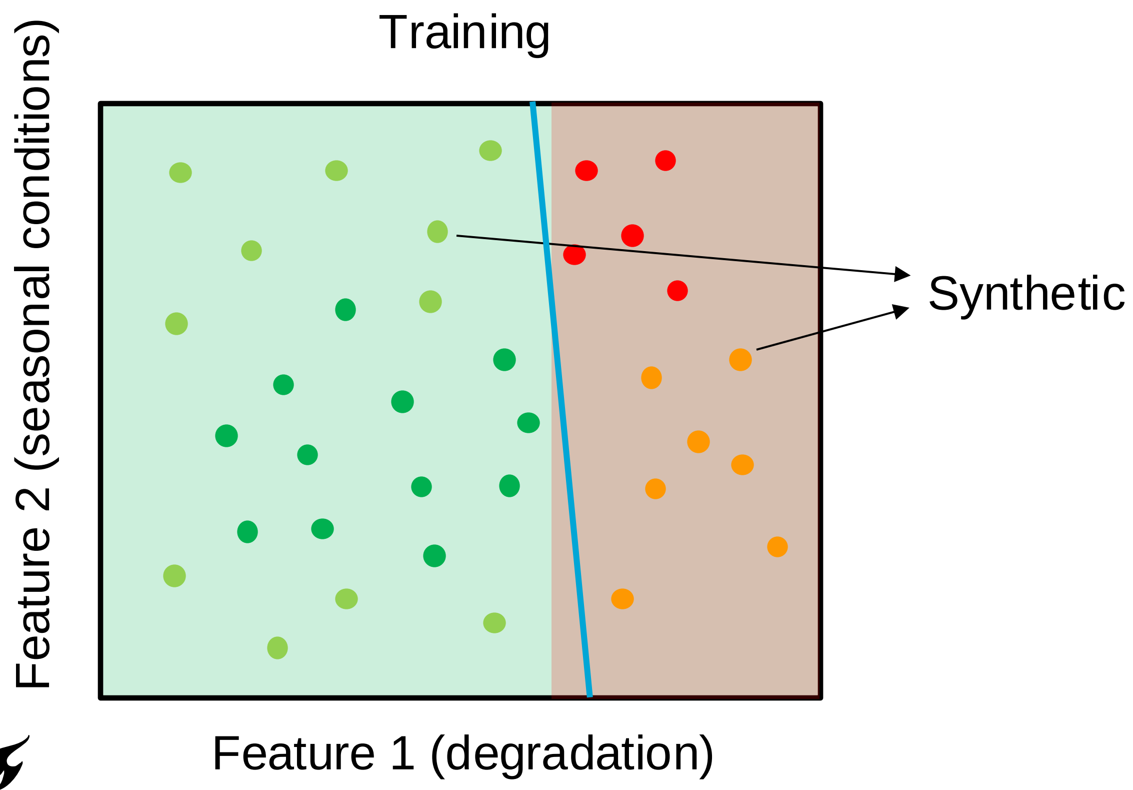
A simple example



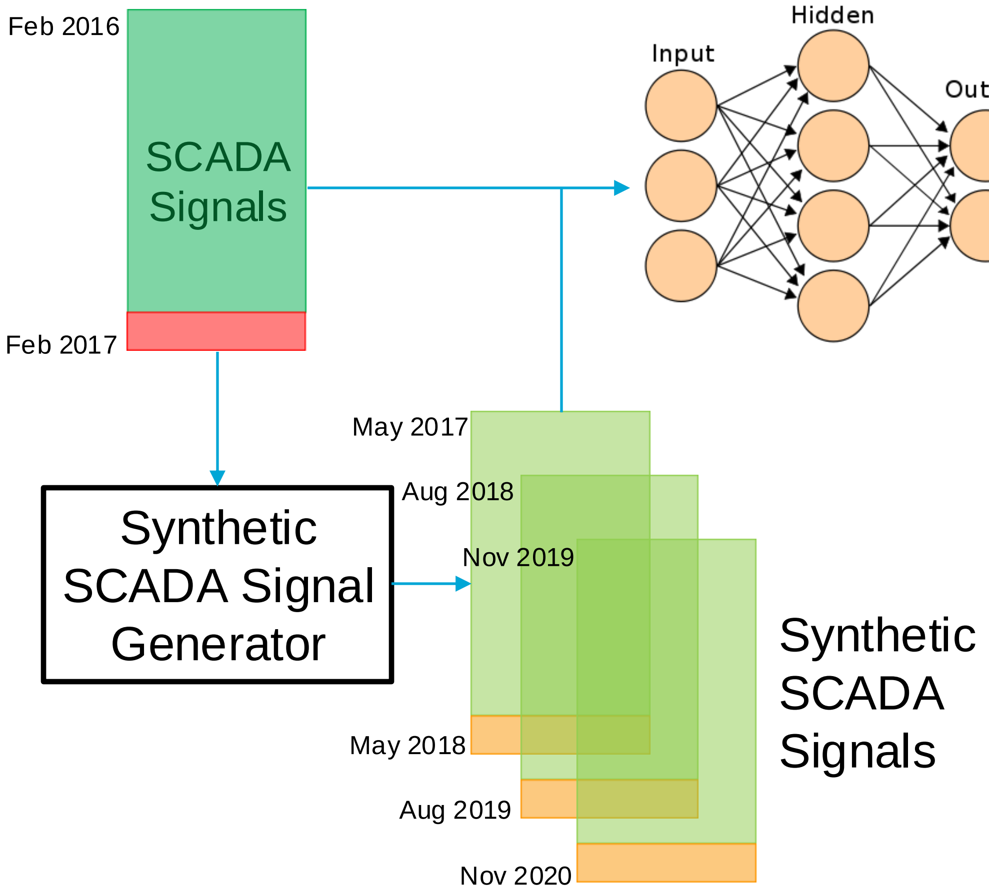
TU Delft

Proposed solution: Synthetic signal generation

Simulating the same failure at different seasonal conditions



<u>Train with a combination of field and synthetic signals</u>



Methodology for synthetic SCADA signal generation

Selected SCADA Signals:

- Related to the operation of the component of interest (e.g., Gearbox temperature and pressure signals) — $\boldsymbol{S}$
- Related to environmental conditions (e.g., Ambient temperature) — $\boldsymbol{E}$
- Related to wind turbine operational mode (e.g., Rotor speed) — $\boldsymbol{O}$

CAE[1]

Health Indicator (HI) — D

Assumption: S is a stochastic function of O, E, and D

$$\boldsymbol{S} = F(\underbrace{\boldsymbol{O}, \boldsymbol{E}}_{\boldsymbol{C}}, D, \boldsymbol{Z})$$

- Signals ← $\boldsymbol{S}$
- Conditions ← $\boldsymbol{C}$
- $\boldsymbol{Z}$ → Unit Gaussian noise vector

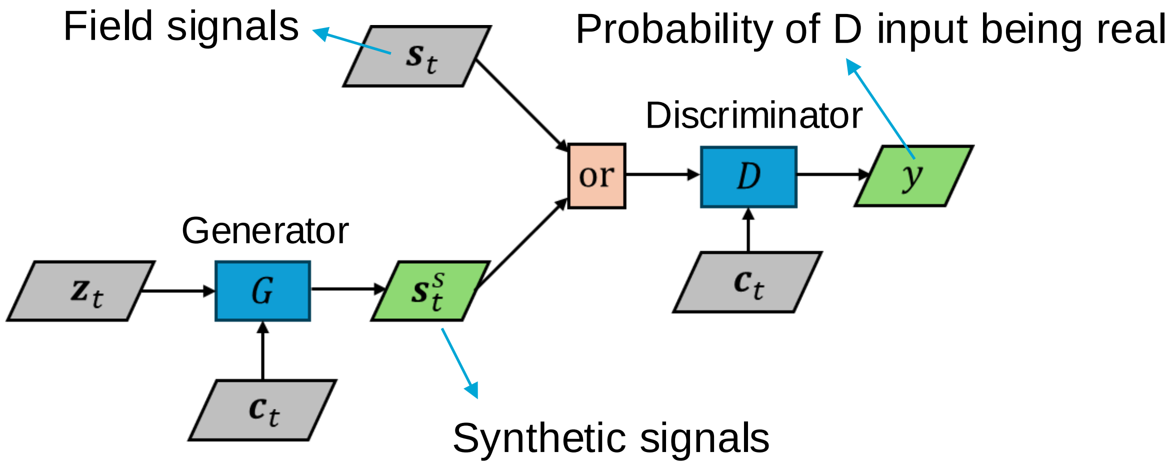
1: Eftekhari Milani, A., Zappalá, D., and Watson, S.; A hybrid Convolutional Autoencoder training algorithm for unsupervised bearing health indicator construction, Engineering Applications of Artificial Intelligence, 2024

Methodology for synthetic SCADA signal generation

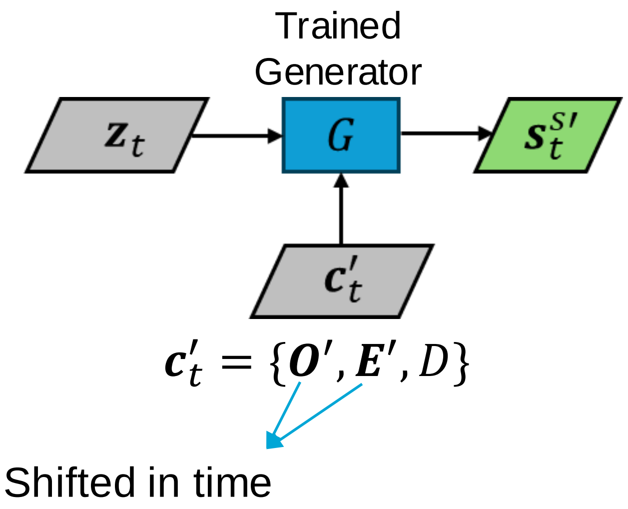
$$\boldsymbol{S} = F(\overbrace{\boldsymbol{O},\boldsymbol{E},D}^{\boldsymbol{C}},\boldsymbol{Z})$$

→ modelled through a Conditional GAN (cGAN) which is a probabilistic generative deep learning framework

Training



Generation



TU Delft

Dataset

- The SCADA dataset is received from **Lucky Wind S.p.A**
- Signals available from nine wind turbines (WT 1-9) through 5 years of operation, resampled to 6-hour intervals
- WT 8 experienced a gearbox failure in Feb 2022 → WT 8 used for training
- The remaining eight WTs → used for testing

- Selected signals for analysis:

$\boldsymbol{S}$ Temperature signals: 3 gearbox bearings, gearbox oil, gearbox oil at the inlet
Pressure signals: gearbox oil before filter, and after filter

$\boldsymbol{E}$ ambient temperature, ambient wind speed

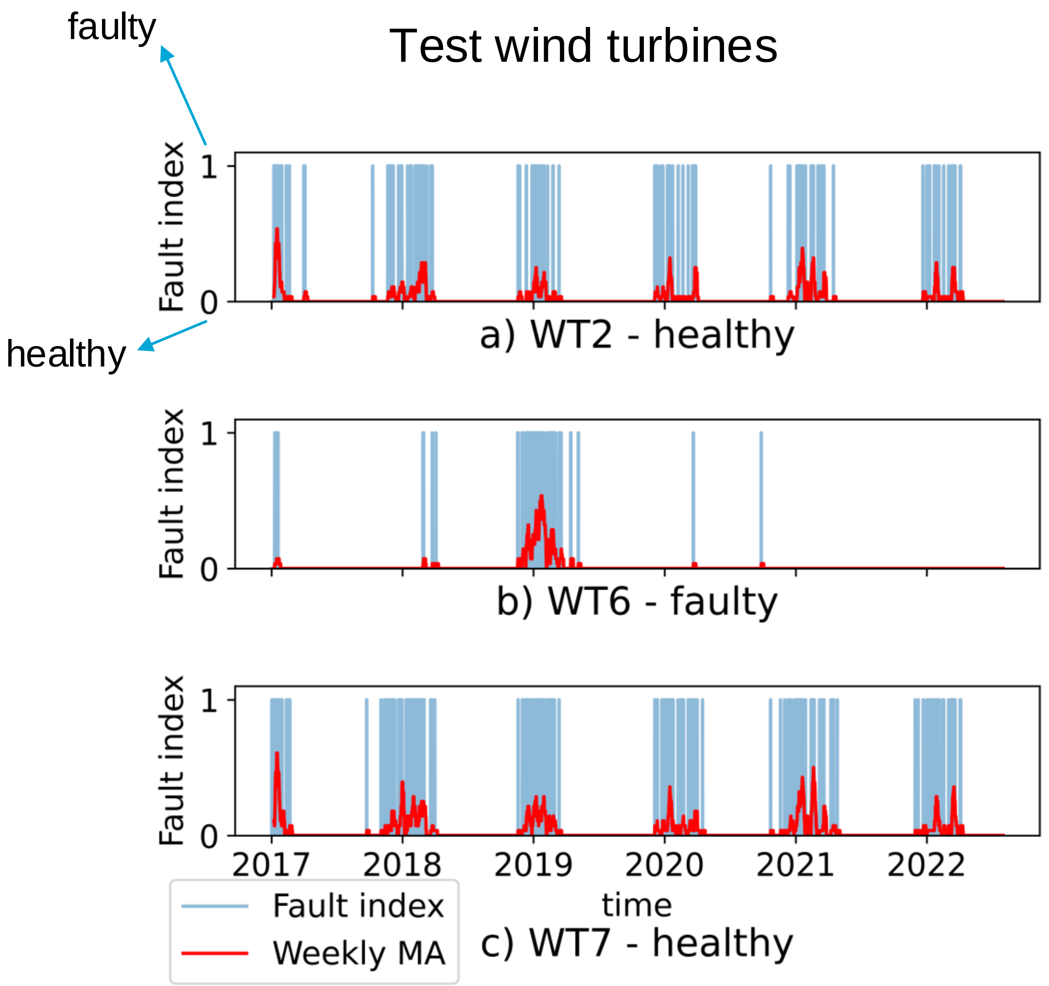
$\boldsymbol{O}$ rotor RPM

TU Delft

Results – Without synthetic signals

ANN trained with data related to only one failure event (WT 8 – one year leading to failure)

- A lot of false positives around the date of the training set failure (February)
- Not clear whether the detection in WT 6 is, in fact, a fault or a false positive



The model cannot distinguish between the features specific to seasonal conditions at the time of failure and those specific to component degradation!

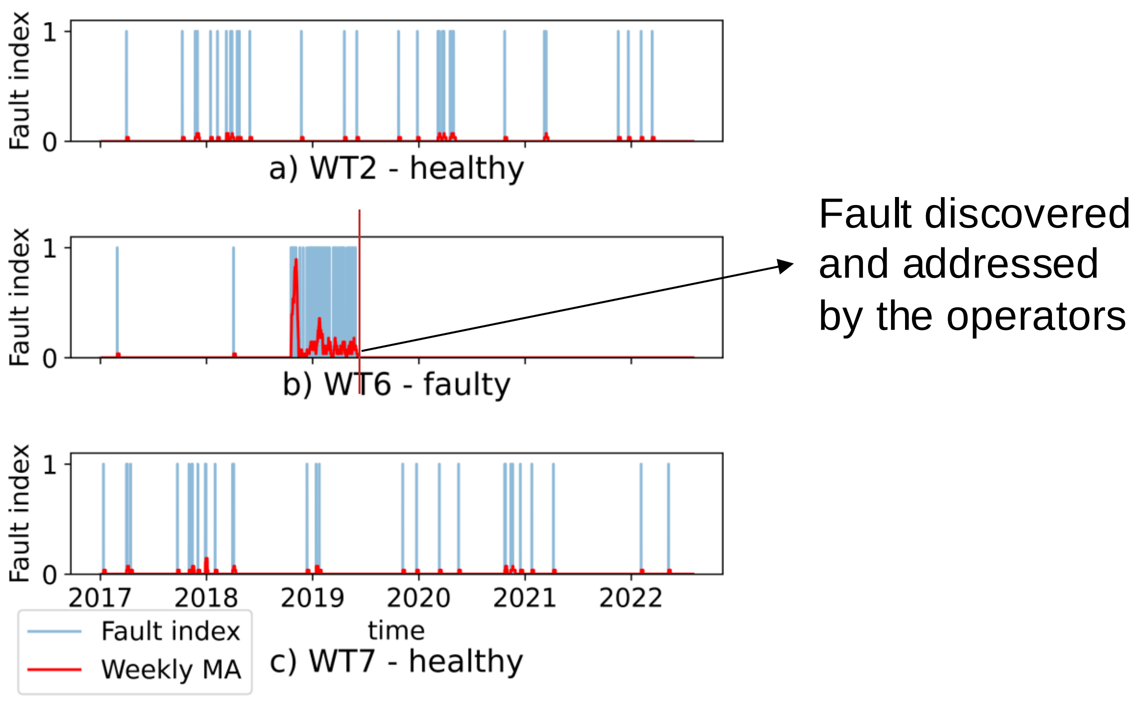
TUDelft

Results – With synthetic signals

ANN trained with both WT 8 data and four synthetically generated data

- Most of the false positives have been resolved
- A fault is detected at the end of Oct 2018 and continues until the end of May 2019
- The operating company confirmed that a gearbox fault was discovered and repaired around the end of May 2019

ANN (Classifier) trained with the field and synthetic datasets



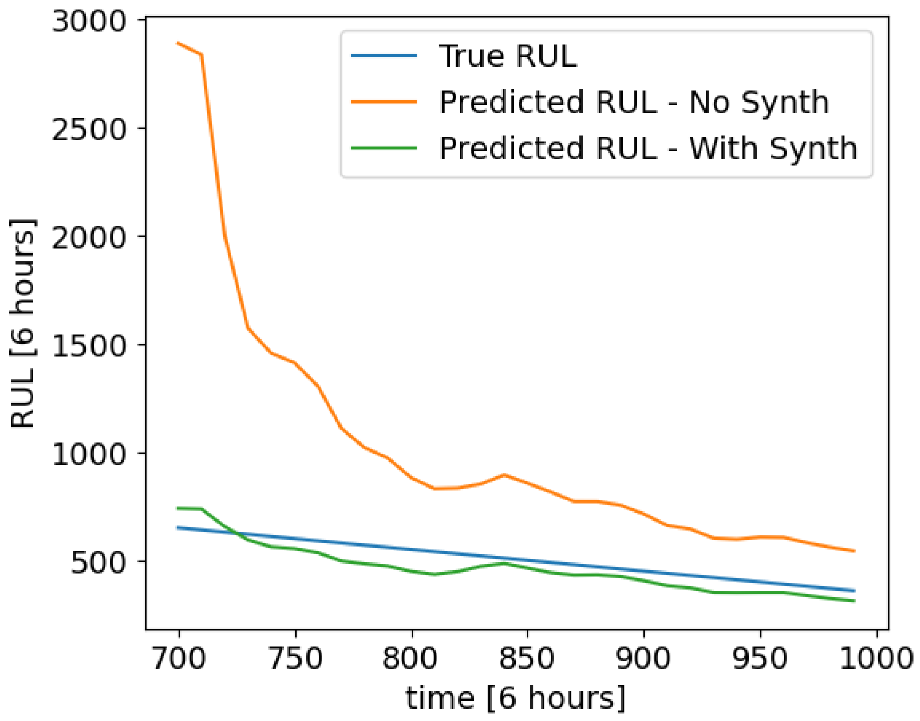
TUDelft

Conclusions

- Classification methods are useful when the detection of a certain type of fault is required. However, their application is hindered by the limited availability of failure events in SCADA datasets
- In this work, a method based on cGANs is proposed to generate synthetic signals and address this limitation
- The generated synthetic signals improve the training of a classifier to learn fault features, leading to significantly fewer false positives.
- This leads to the blind detection of a fault in the gearbox of another wind turbine seven months before its discovery and maintenance by the wind farm operators.

Ongoing work

- Application to Remaining Useful Life (RUL) prediction



TU Delft

Thank you for your attention!