CNNs for Fault Detection: a Wind Turbine Use Case

Lilach Goren Huber
Markus Ulmer
Jannik Zgraggen

lilach.gorenhuber@zhaw.ch

BBS series

1) CNNs for Scalable Fault Detection: a Wind Turbine Use Case (Lilach, 29.4.2021).

2) Transfer Learning Approaches for Fault Detection (Jannik, 6.5.2021).

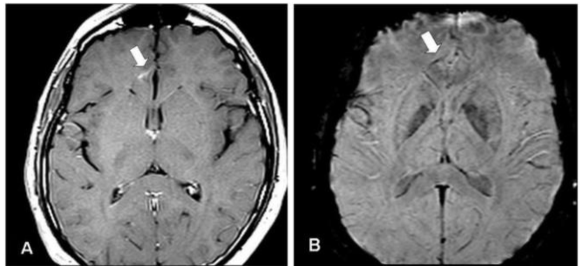
Lilach Goren Huber

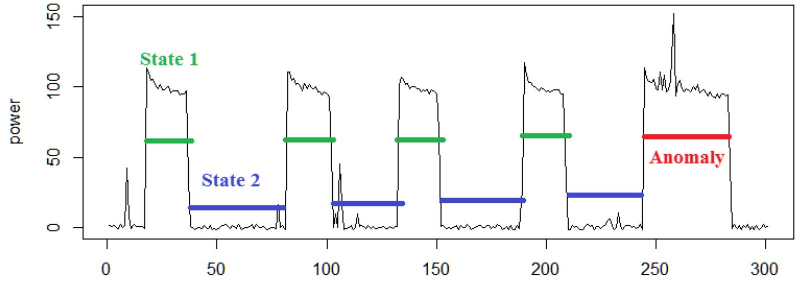
Markus Ulmer

Jannik Zgraggen

Fault Detection

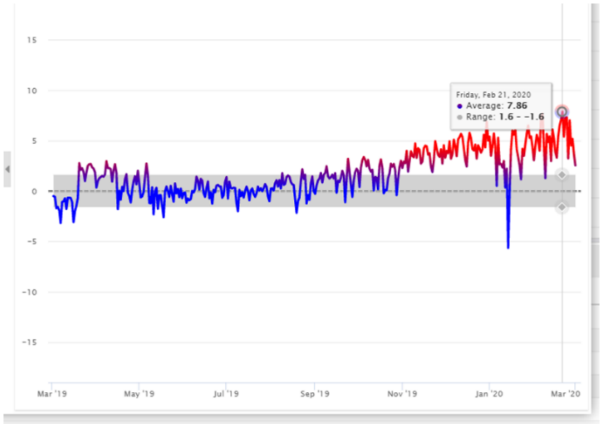
- Anomaly detection: detect deviations from a familiar pattern ➔ **anomaly score**.
- Fault detection: detect early patterns of deviations from normal behavior of machines ➔ **Health Index**.
- Common machine data: multivariate time series.





Complexities

- Anomalies are **rare** ➔ class imbalance
- **Unknown** fault nature (time dependence, distribution)
- **Diversity** of anomaly classes/types (abrupt, slow degradations, point anomalies, different severity)



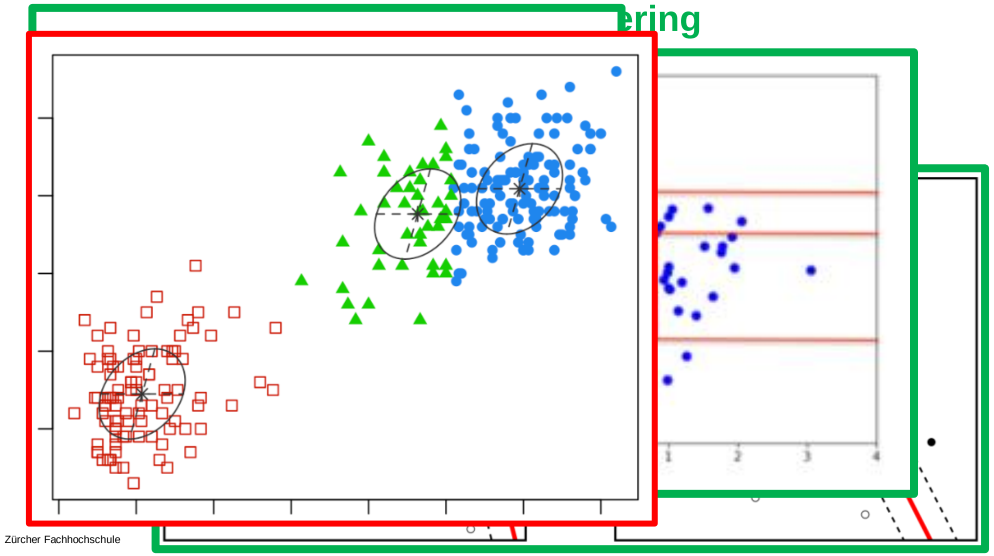
Challenges

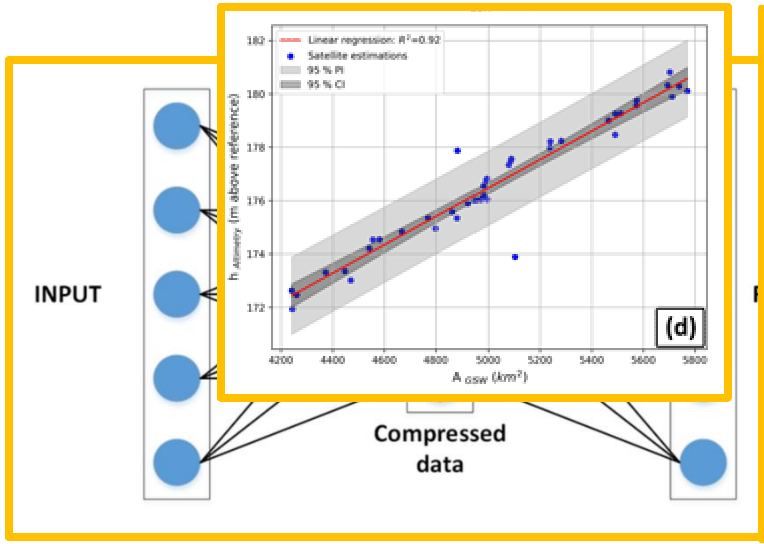
- Thershold setting ➔ False positives vs. Missed detections.
- Model evaluation.
- Detection + Explanation => diagnostics.
- Noise resilience under diverse operating conditions.

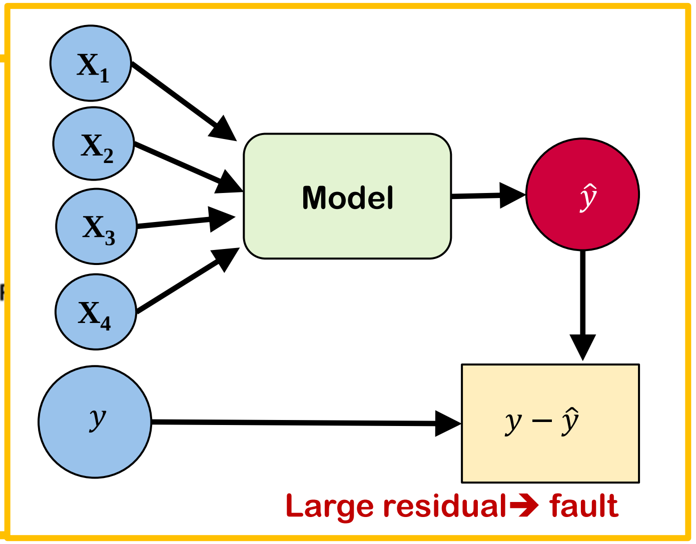
Anomaly Detection

Need labeled faults | **Few clear outliers** | **Train on healthy data**

- **Supervised** (Need labeled faults)
  - Classification
- **Unsupervised** (Few clear outliers)
  - Distance/density based
    - Clustering
    - OCSVM
    - Isolation forest
- **Semi-supervised** (Train on healthy data)
  - Dimension Reduction
    - PCA
    - NNs
  - Reconstruction
    - PCA
    - Autoencoder
  - Regression/prediction
    - Stat. models
    - AR models
    - Reg. Trees
    - NNs



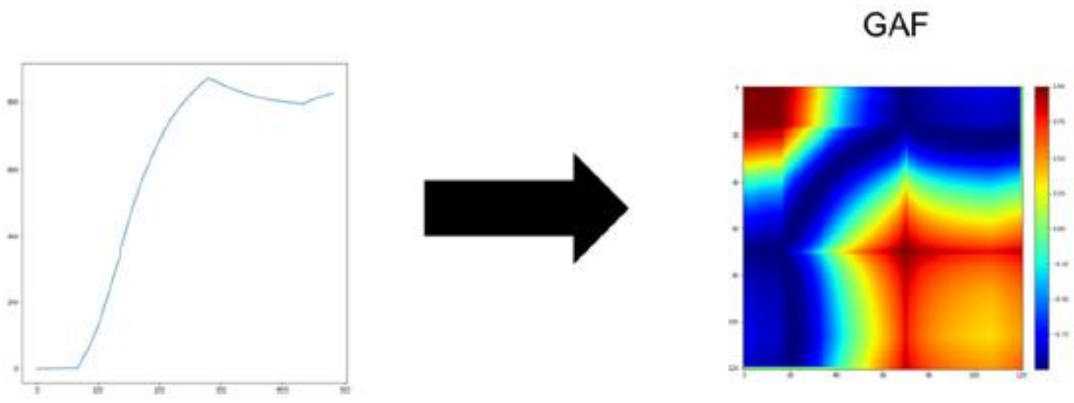


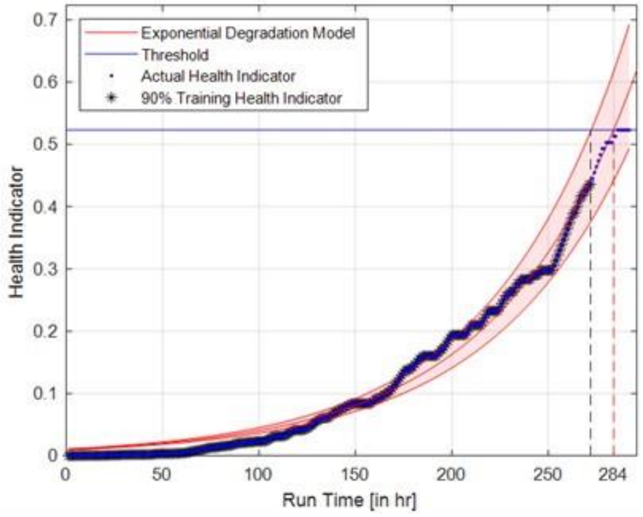


Trends in Fault Detection

Academic Research

- Method driven ➔ classification mainly
- Encode ts data as images ➔ limited applicability.
- RUL predictions (simulated data)
- Transfer Learning/ Domain Adaptation
- Hybrid models (physics + ML)
- Interpretability («XAI»)





Trends in Fault Detection

Academic Research

- Method driven ➔ classification mainly
- Encode ts data as images ➔ limited applicability.
- RUL predictions (simulated data)
- Transfer Learning/ Domain Adaptation
- Hybrid models (physics + ML)
- Interpretability («XAI»)

Practical R&D

- Use existing data (little, no faults, low freq)
- How to set robust thresholds?
- How to evaluate algorithms without labels?
- How to scale up algorithms (data selection, run times)?
- Robustness under changing conditions.
- Diverse operating conditions (TL/DA?)

Who are we?

Zurich University of Applied Sciences
School of Engineering

Smart Maintenance Team

Nispera®
renewable energy data solutions

Data intelligence platform for renewable energy assets

> 7 GW of monitored assets

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Innosuisse – Swiss Innovation Agency

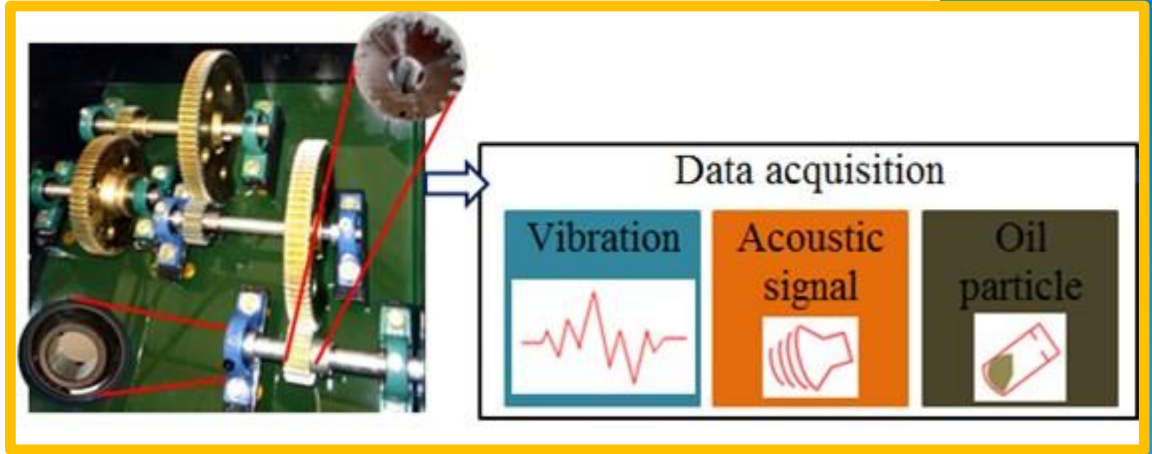
EKZ

Electricity provider of Canton of Zurich

Intelligent fault detection and diagnosis algorithms for wind turbines

https://nispera.com/solutions/predictive-maintenance

Added Value for Park Operators



Original Equipment Manufacturer (OEM):

- Designated Condition Monitoring hardware.
- High frequency data.
- Use for design improvement.

Not accessible for operators/owners.

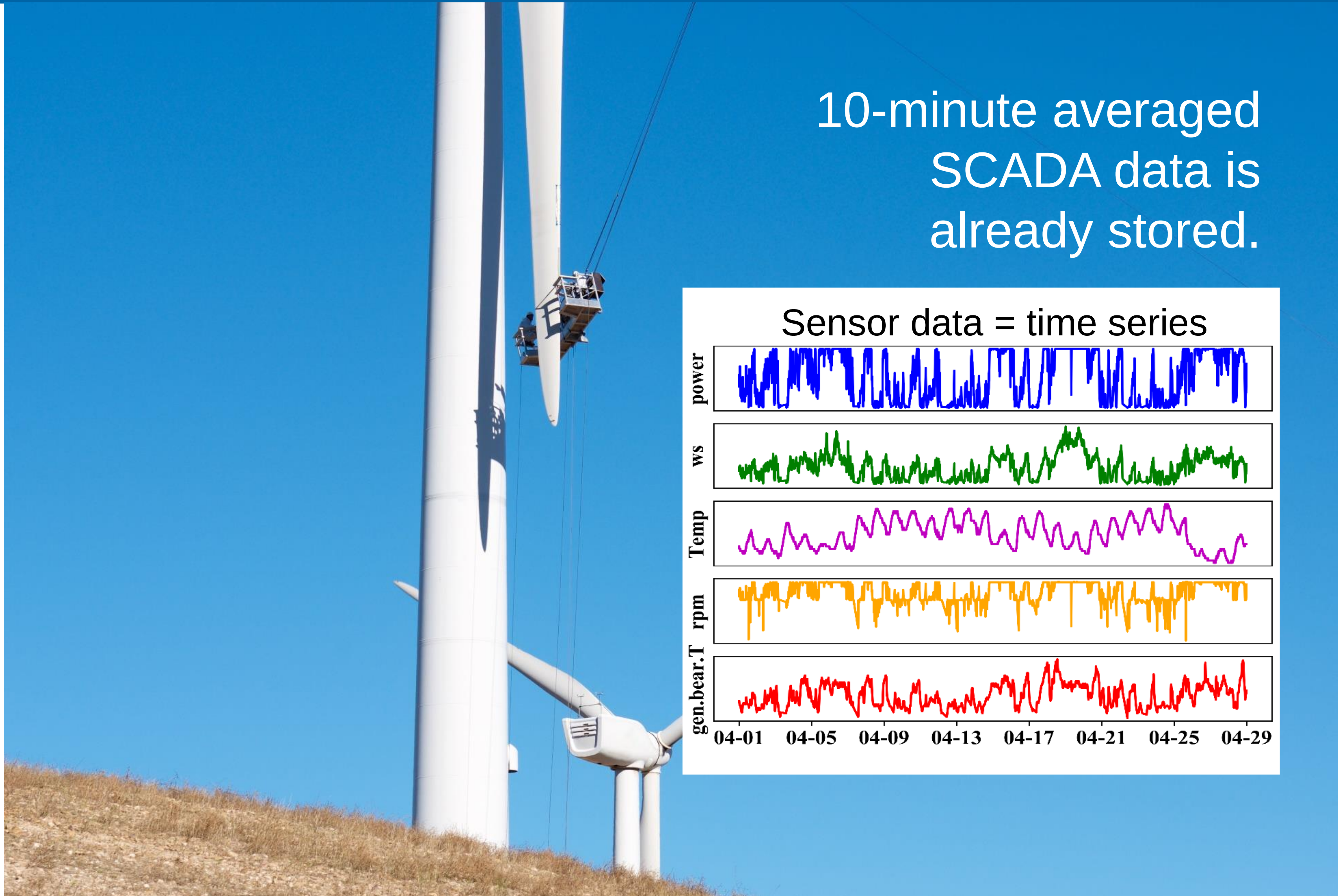
Designated Condition Monitoring Systems are too expensive

Added Value for Park Operators

Zurich University of Applied Sciences
zh aw School of Engineering

Cost effective solution:

- Use available data.
- No need for additional hardware.



10-minute averaged SCADA data is already stored.

Challenge
High performance fault detection with existing data

Zurich University of Applied Sciences
zh aw School of Engineering

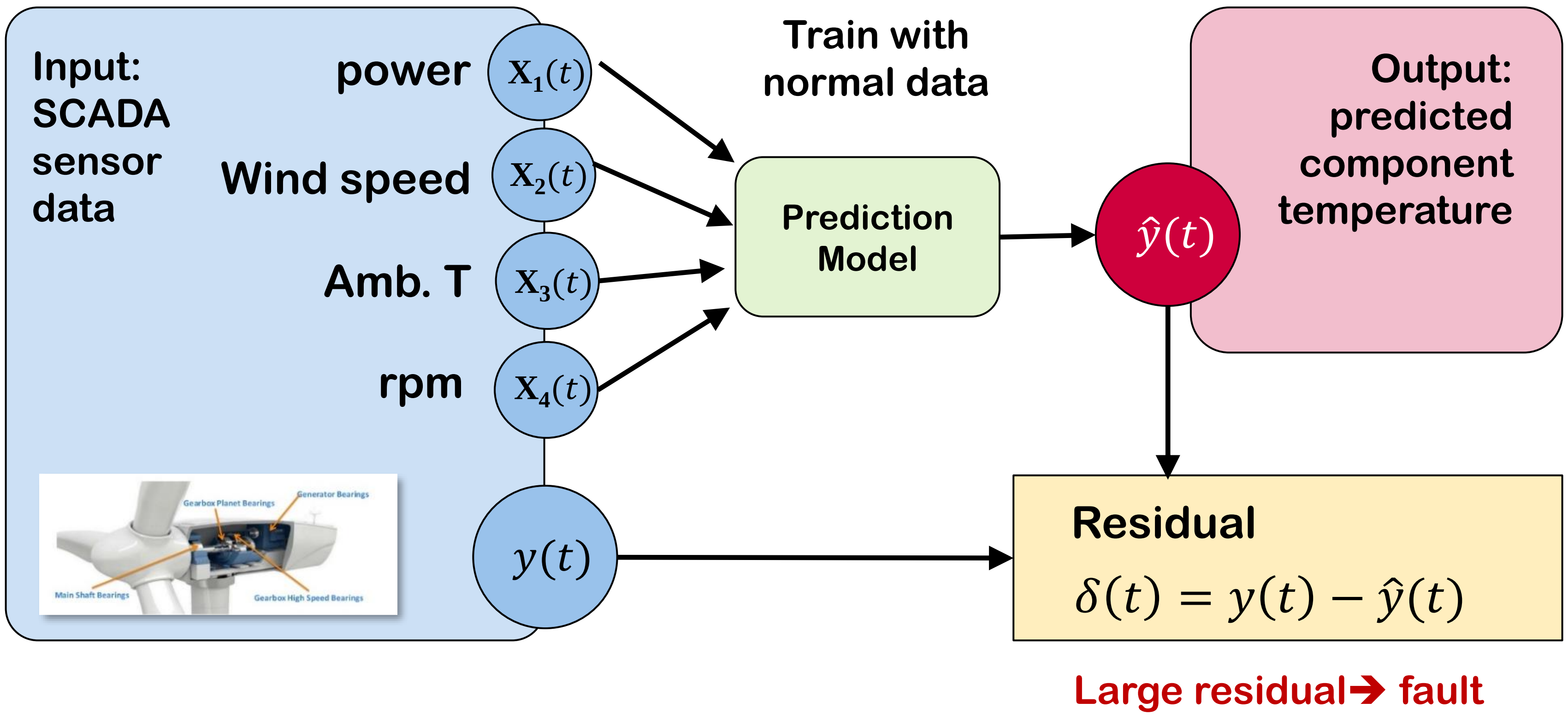
10-minute averaged SCADA data is already stored.

Practical solution:

- Detect early
- Accurately
- Diverse Fault types
- Transferable
- Scalable

Semi-supervised Anomaly Detection

- ✓Faults are rare and unique ➔ use only healthy data for training, detect anomalies online.
- ✓Fault localization ➔ regression



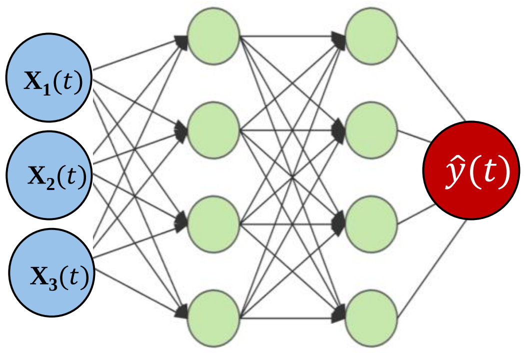
Prediction model
Single Output CNN

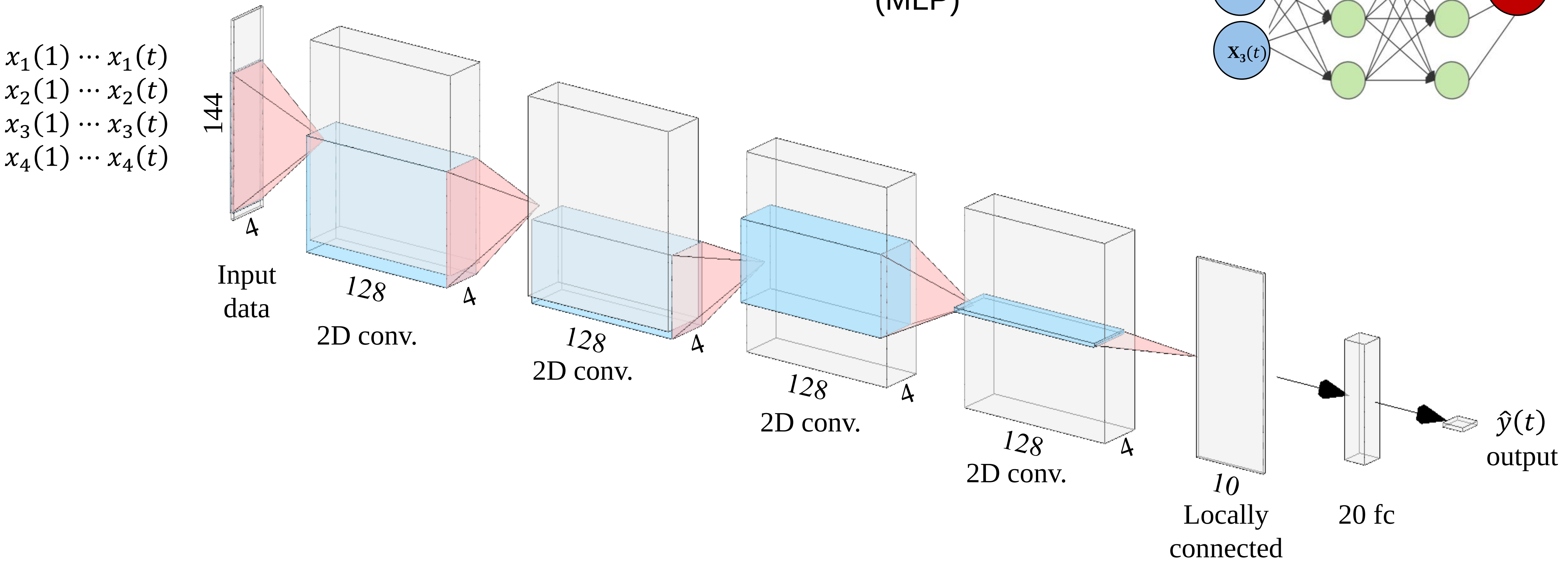
Zurich University of Applied Sciences

zh aw School of Engineering

- ✓ Learn time dependent patterns ➔ CNN

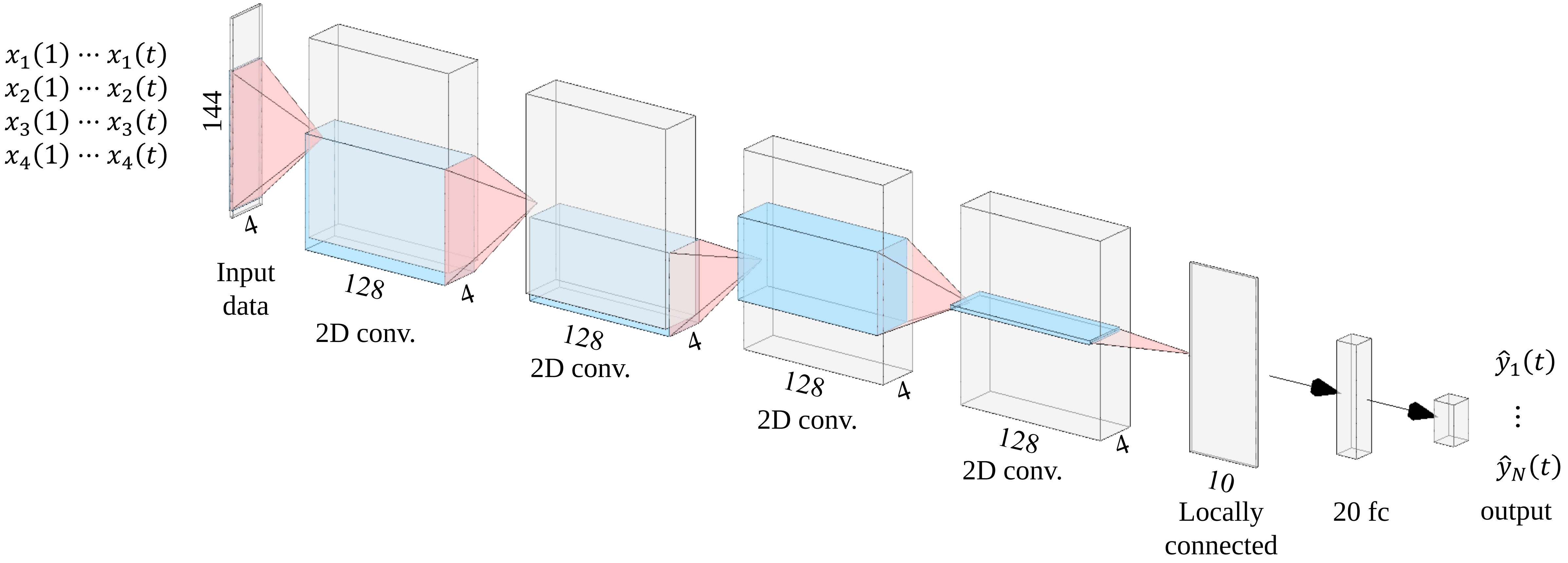
Commonly used:
Multi Layer Perceptron (MLP)





Prediction Model
Multi-Output CNN (CNNm)

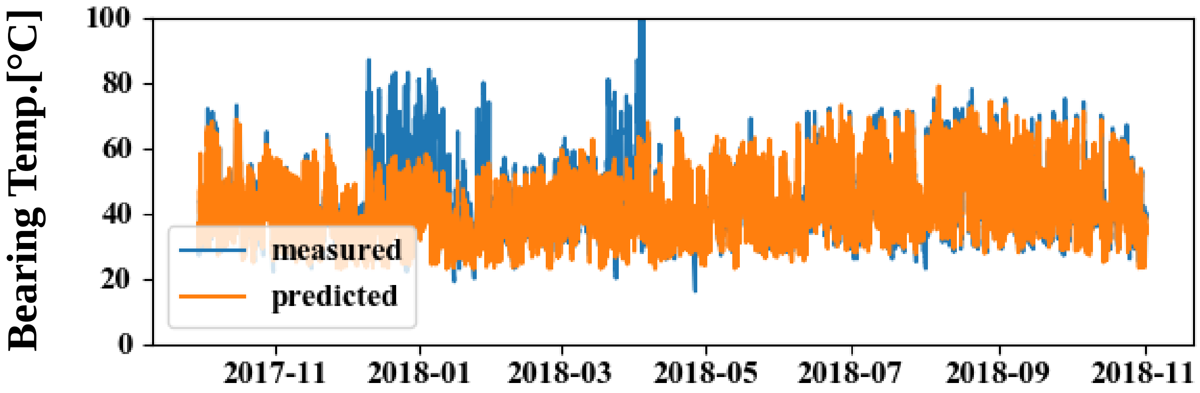
✓ Diverse fault types ➔ multi-target regression



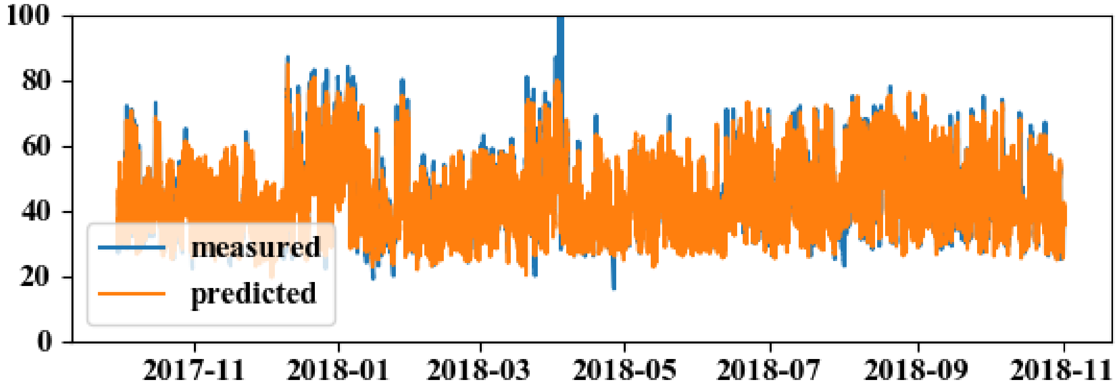
Model Evaluation and Selection

- Regression models: metric = prediction error MSE.
- AD: training on healthy data ➔ we can only evaluate prediction error and minimize it.
- **Problem: model selection.** reducing the error on the healthy data does not imply increased error for fault data.
- Example: we can select a predictor with perfect correlation to the target. We then have good predictions also during abnormalities.

Without nacelle T as input

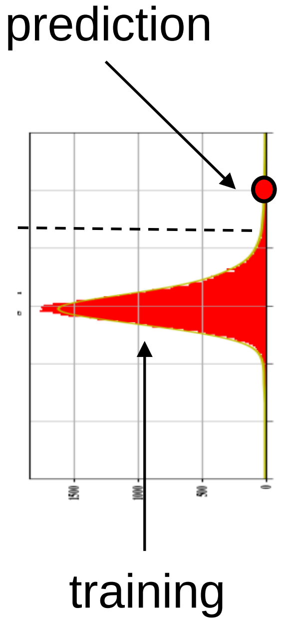


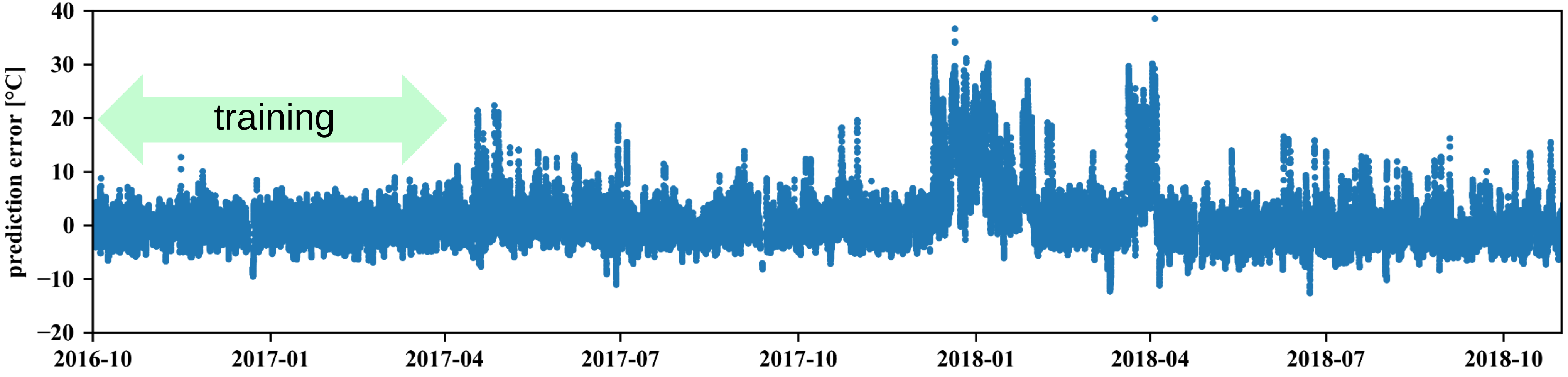
With nacelle T as input



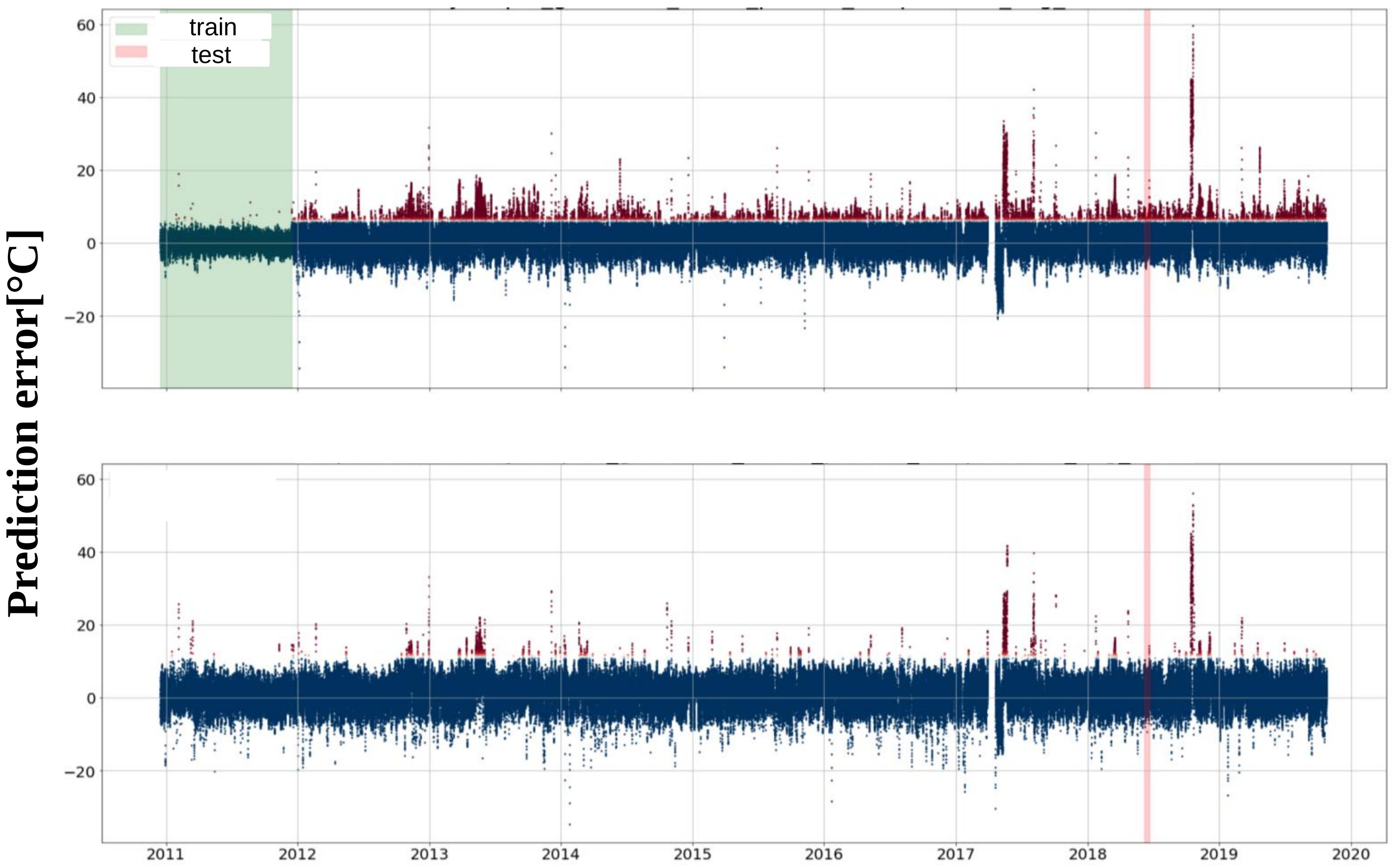
➔ select a minimal set of «exogenic» predictors.

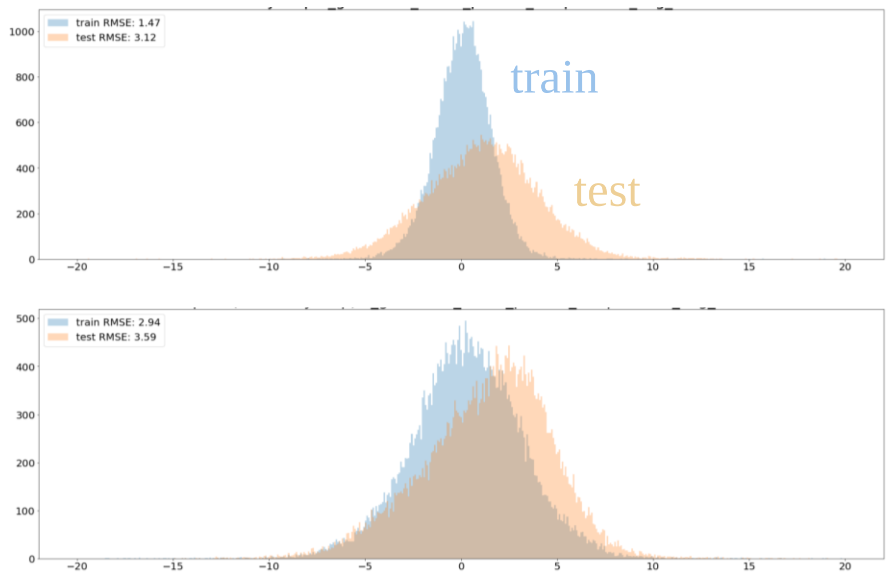
Over fitting





Over fitting





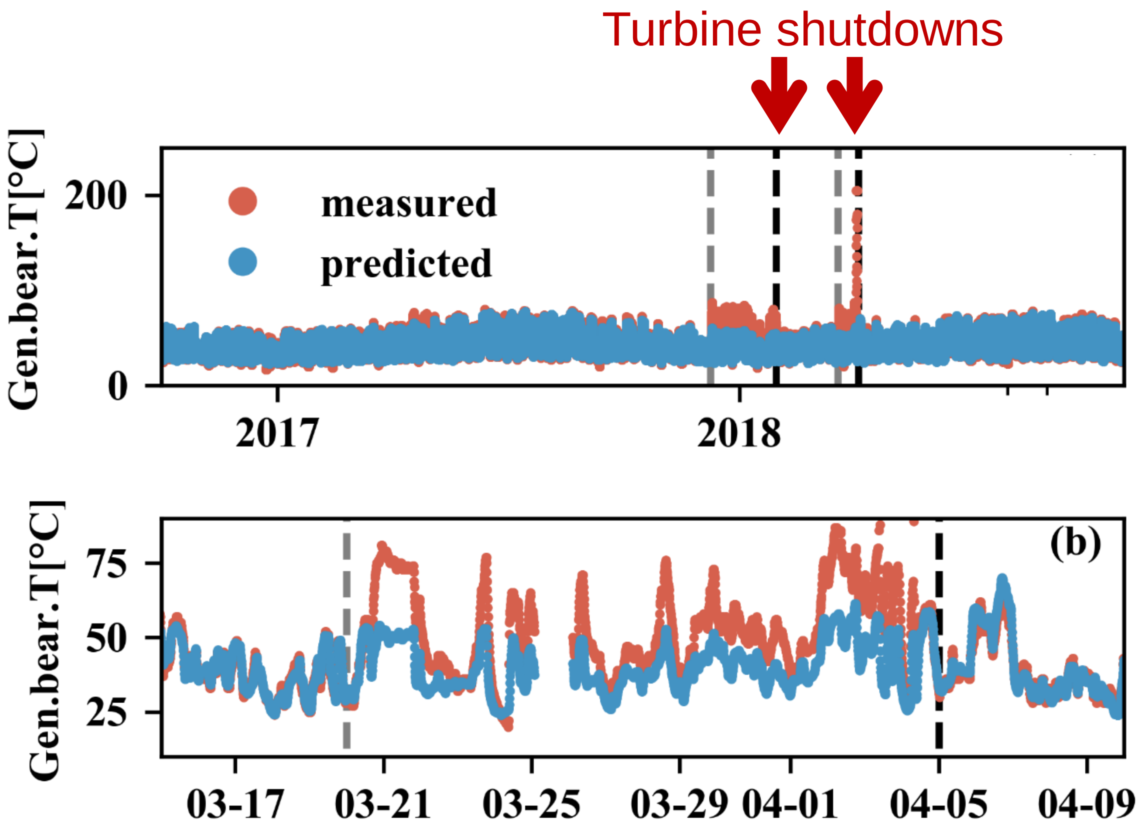
Over fitting:
Prediction errors at testing do not grow
training get smaller

➔ False positives.

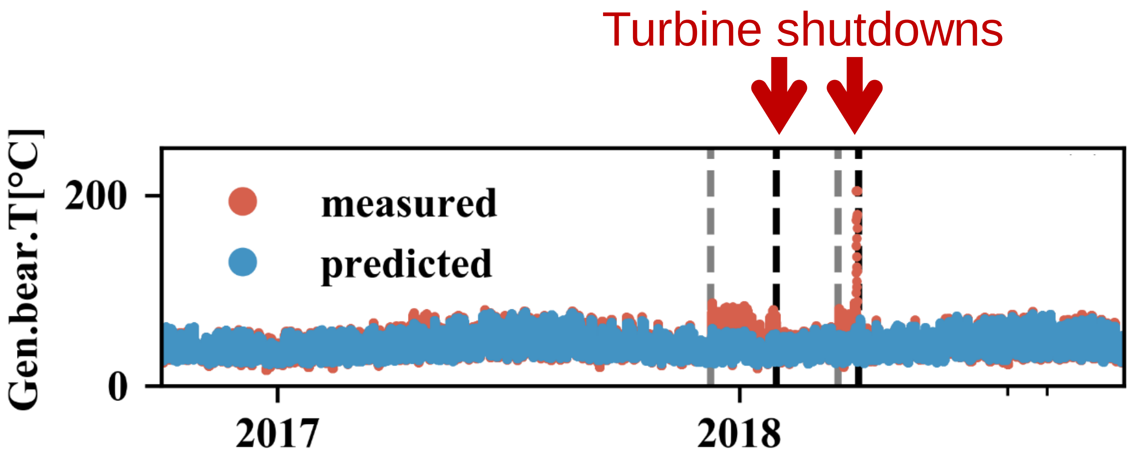
Example
Detection of Abrupt Faults

Turbine A_0

Abrupt faults in the generator bearing



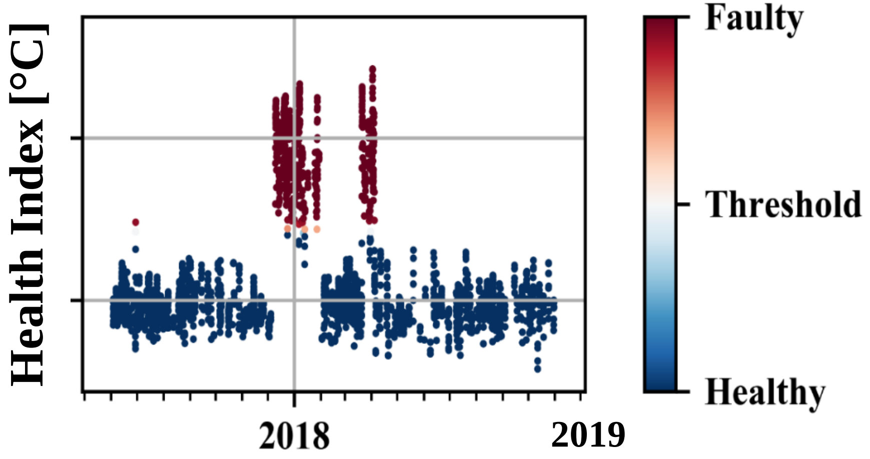
Results
Residuals and Health Index: Abrupt Faults

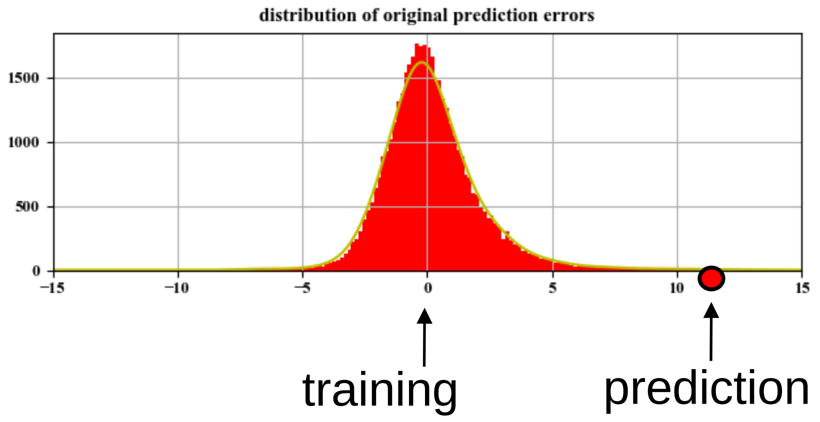


Turbine A_0

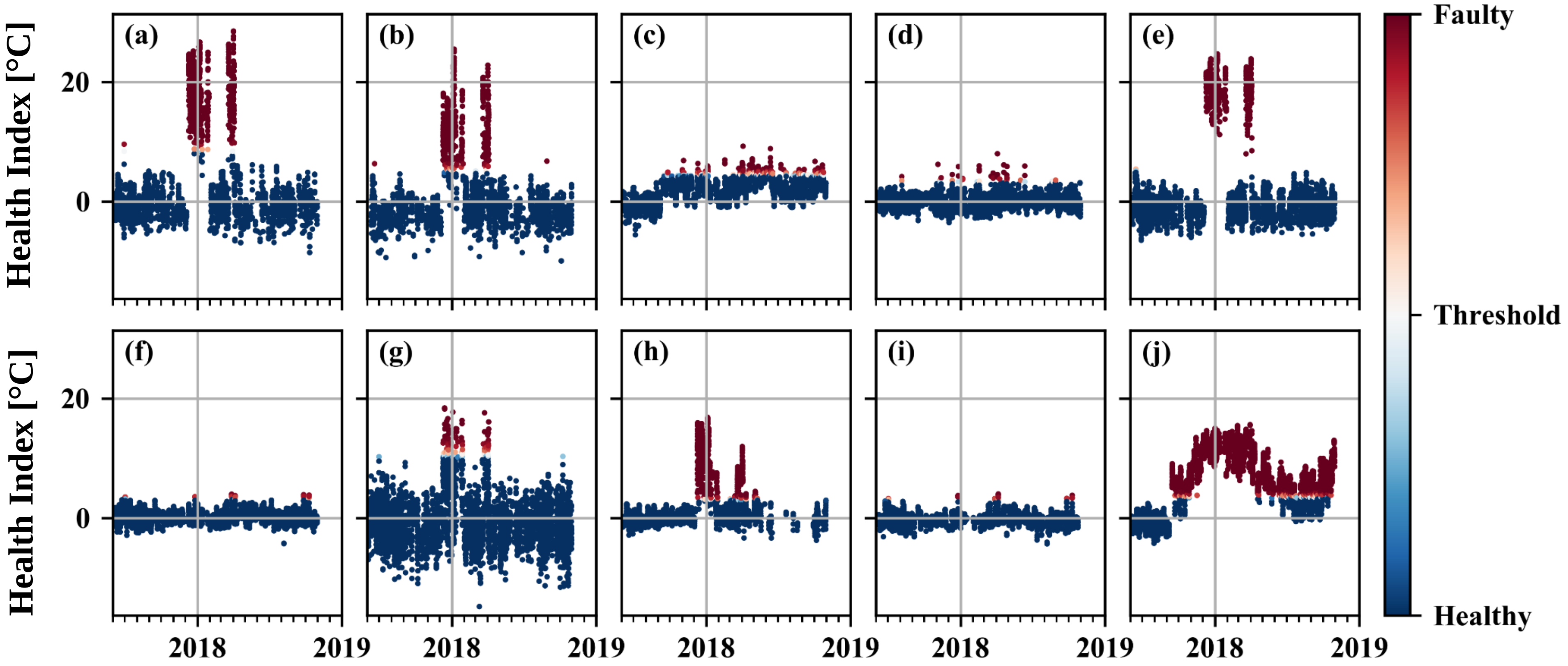
Abrupt faults in the generator bearing

$$\delta(t) = y(t) - \hat{y}(t)$$ ➡ Health Index



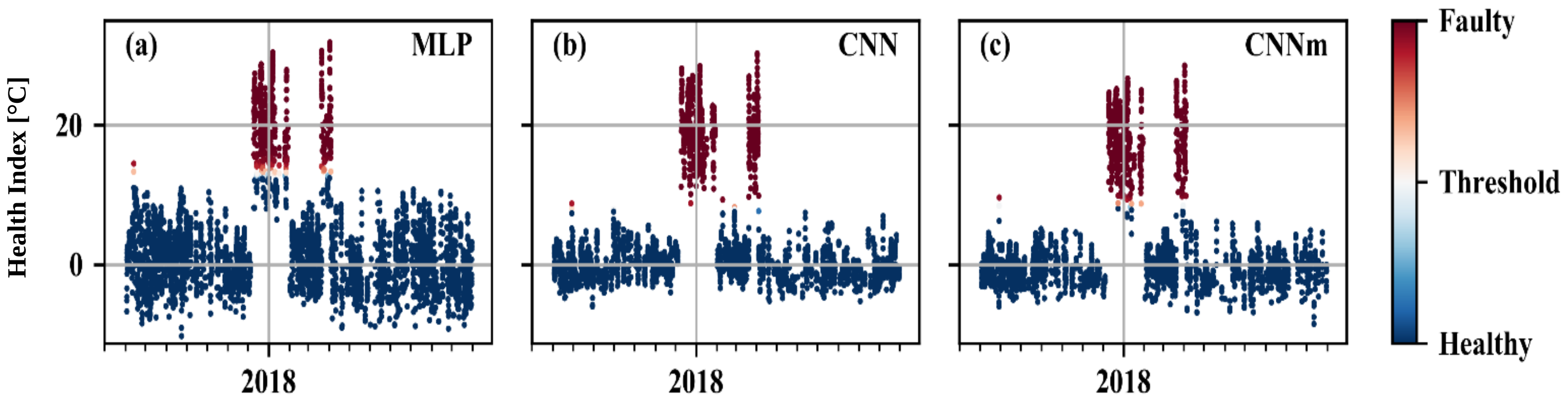


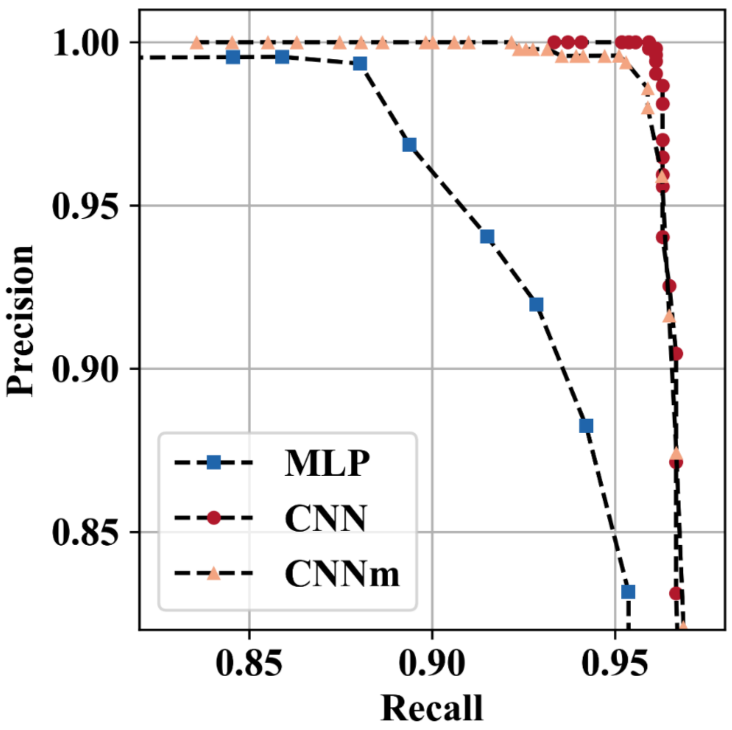
Results
Health Indices with Multi-output CNN



(a) Generator bearing (b) stator phase (c) gearbox bearing (d) gearbox oil (e) slip ring (f) rotor spinner (g) grid transformer (h) controller top (i) controller hub (j) hydraulic oil.

Results
Model Comparison: Abrupt Faults



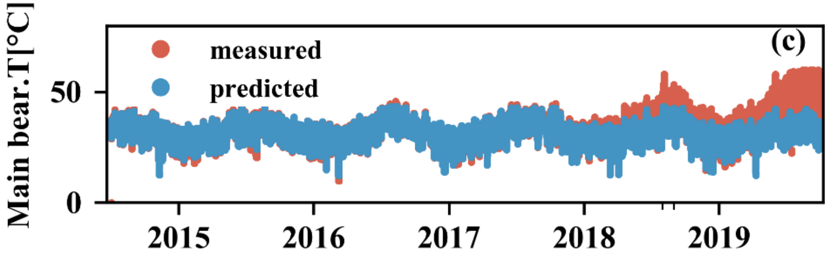


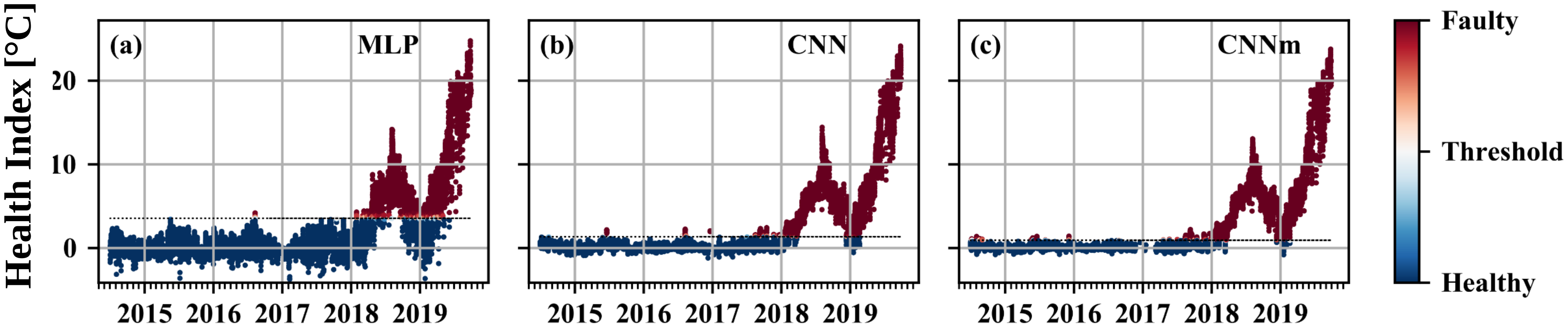
Both CNNs perform better than the MLP

Results
Model Comparison: Slow Degradation

Turbine B_0

Slow degradation of the main bearing

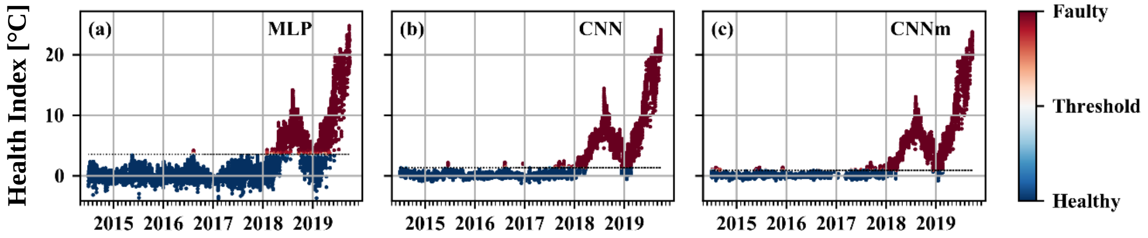




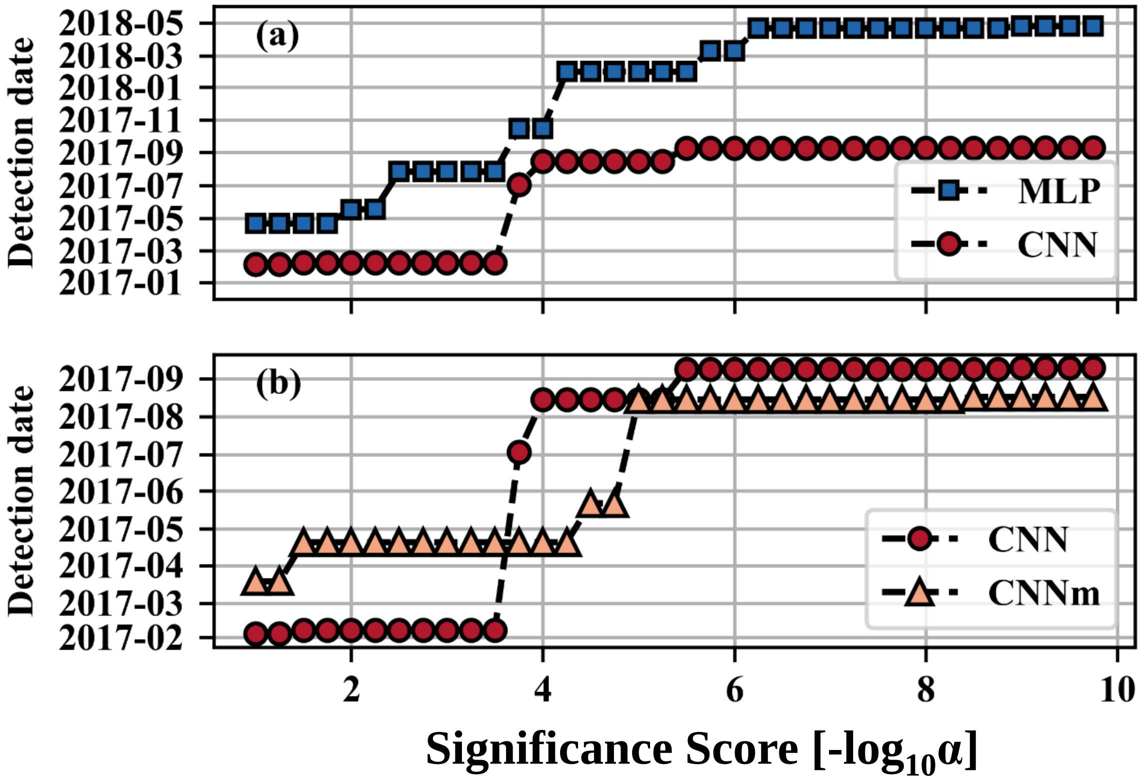
Model	First detection (α=0.0001)	First detection high confidence
MLP	16.10.17 2:00	25.4.18 00:00
CNN	15.8.17 17:00	10.9.17 19:00
CNNm	15.8.17 17:00	9.9.17 14:00

- ✓ CNN detects 2 months earlier than MLP.
- ✓ CNNm detects even earlier than CNN.

Results
Sensitivity Analysis: Slow Degradation



Modify the detection threshold and calculate the date of first detection:



Ulmer, Markus, et al. "Early fault detection based on wind turbine scada data using convolutional neural networks." *5th European Conference of the Prognostics and Health Management Society,* 2020.

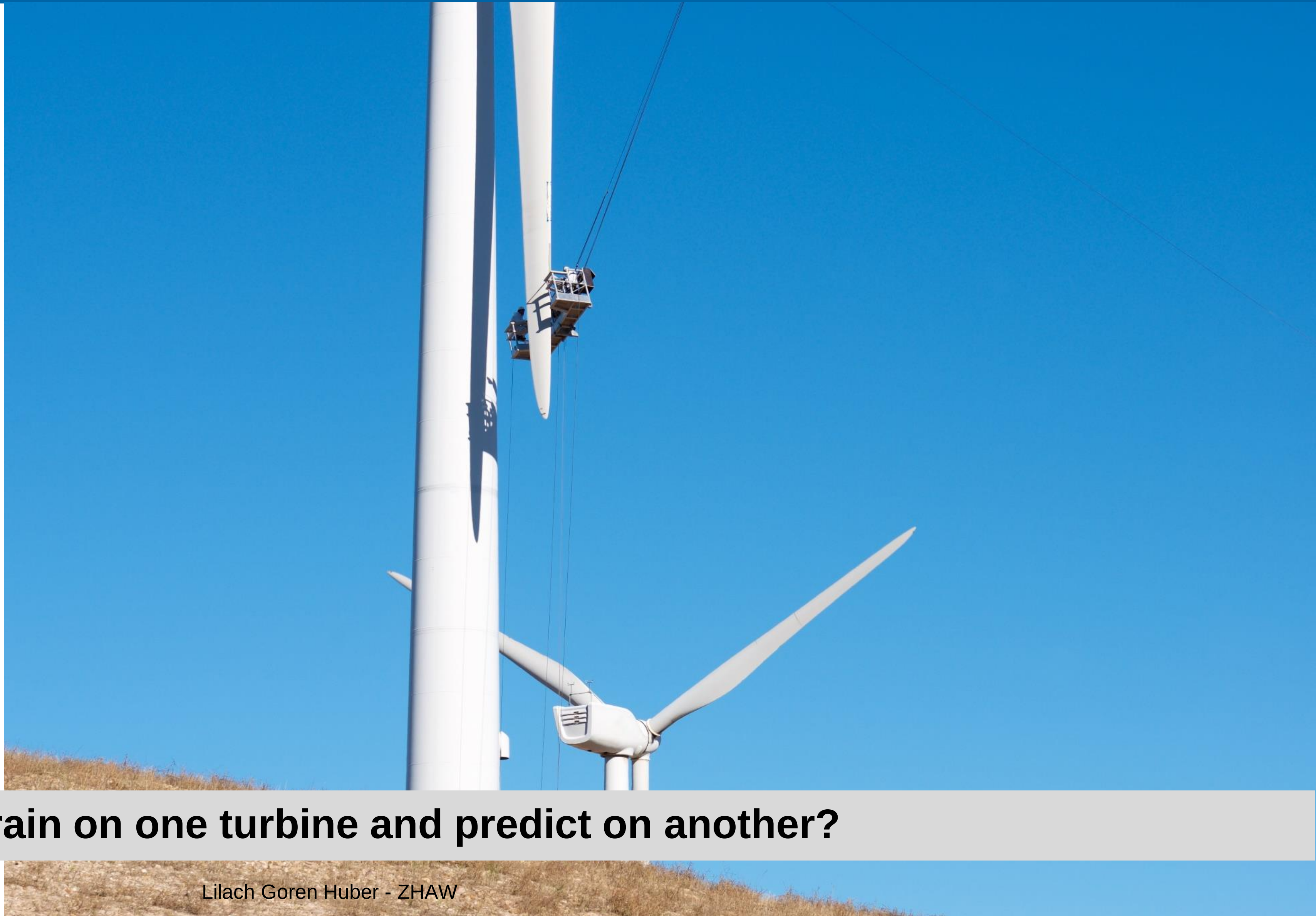
The CNNm is more robust against threshold selection

Scaling up fault detection algorithms

Zurich University of Applied Sciences
zh aw School of Engineering

Motivation for transfer learning

- New turbines/ farms.
- New operational conditions.
- Speed up fleet-wide detection.

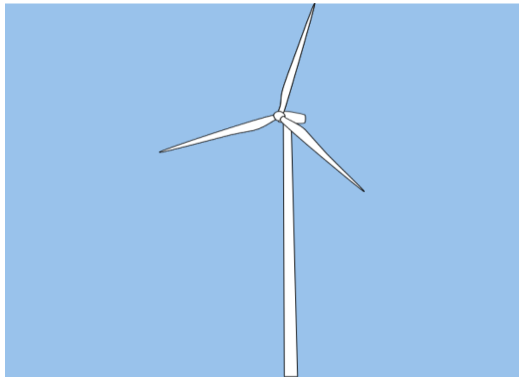


Can we train on one turbine and predict on another?

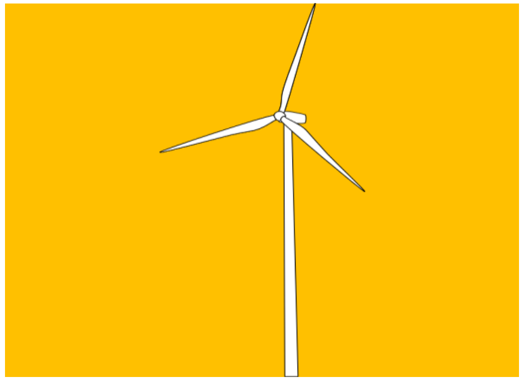
Solution
Cross-Turbine Training Scheme

Goal: predict faults on turbine T.

Problem: only 3M data from T.



Source Turbine S
Available healthy data: 9 Months

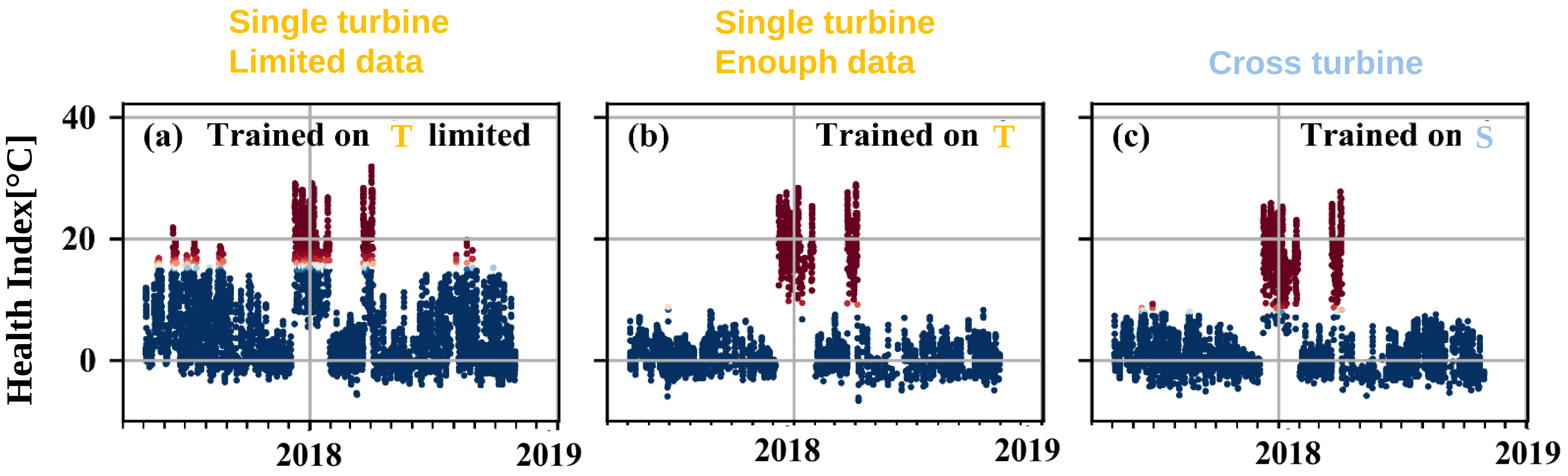


Target Turbine T
Available healthy data: 3 Months

Simple Solution:

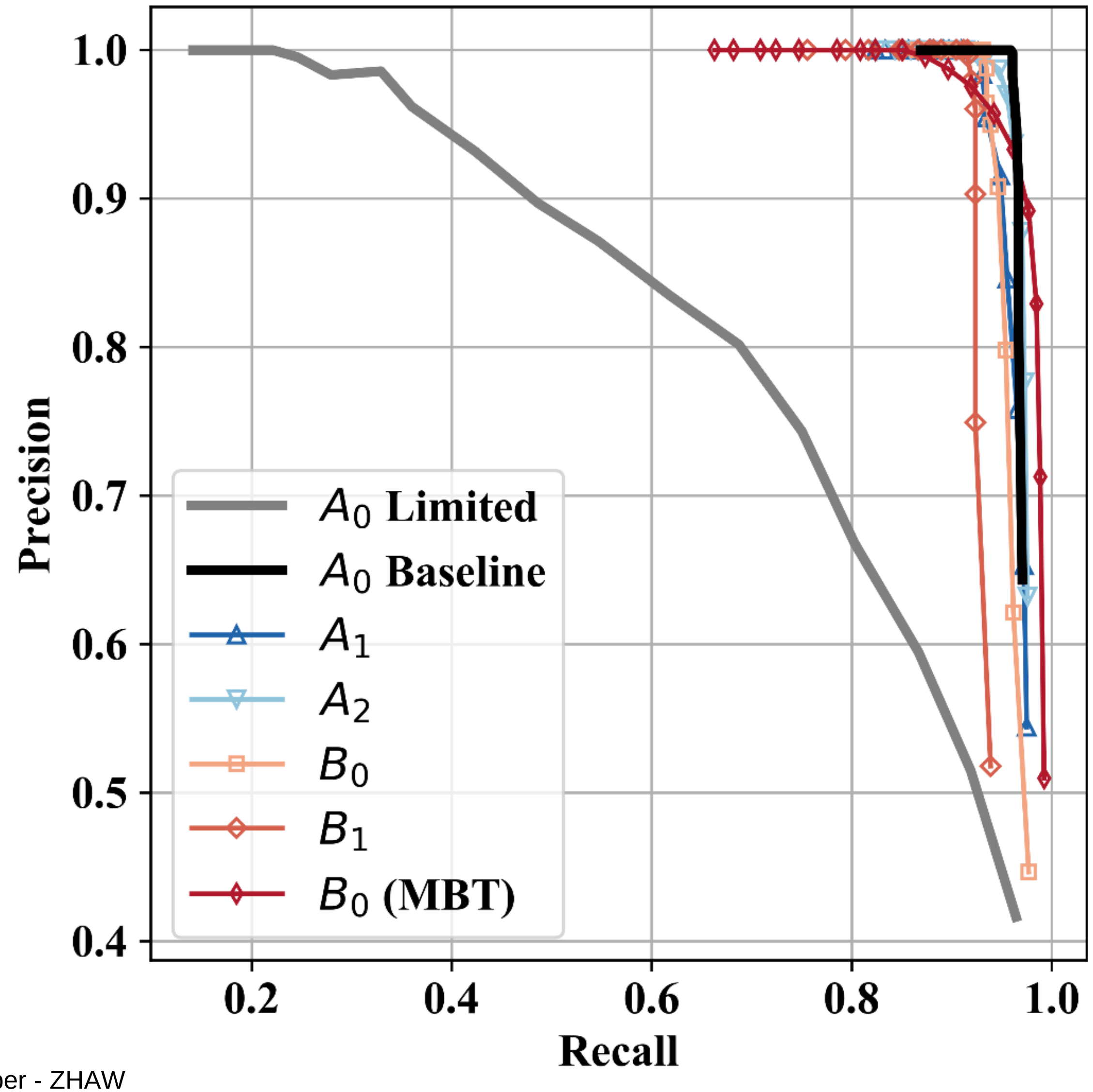
- Train CNN on 9M from turbine S.
- Predict $\hat{y}_t$ for new data from T.
- Train a regression model $y_t \sim \hat{y}_t + X_t$ on 3M from T
- Adapt new predictions on T using the same regression.

Solution
Cross-Turbine Training Scheme



Cross-turbine training scheme allows early fault detection with scarce data

Results
Source domain comparison: abrupt fault



Ulmer, Markus, et al. "Cross-turbine training of convolutional neural networks for SCADA-based fault detection in wind turbines." *Annual Conference of the PHM Society.,* 2020.

Summary

- Fault detection of rare and diverse fault types ➔ semi-supervised AD
- Easy fault localization ➔ regression
- Robust capturing of time depentent patterns ➔ CNNs
- Scalable multi-component detection ➔ multi-output CNN.
- Robust threshold setting ➔ use error distributions for anomaly scores.
- Model evaluation ➔ MSE not enough.

- Scarce data ➔ TL approaches.
- Preliminary: LR-based cross turbine training scheme with promising results.
- Next week: overview of TL approaches, comparison of selected solutions.