Scalable Deep Learning Algorithms for PHM in Operational Fleets

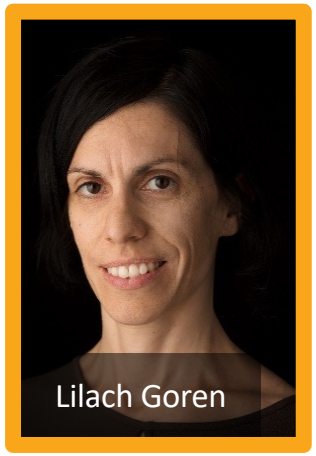
Bridging the research–industry gap

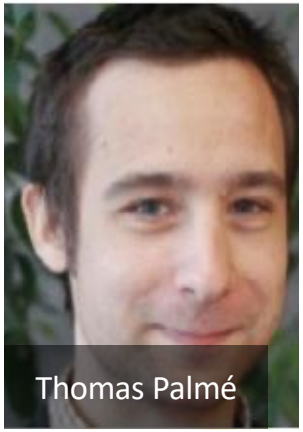
Dr. Lilach Goren Huber
Dr. Manuel Arias Chao
Smart Maintenance Team
Zurich University of Applied Sciences

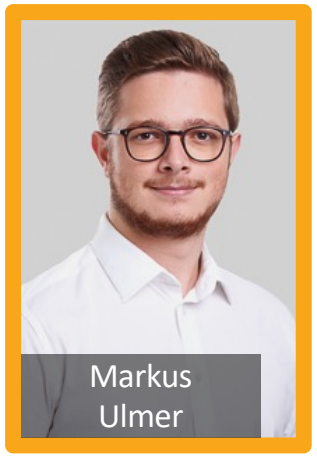
PHM Conference, Nov 2022

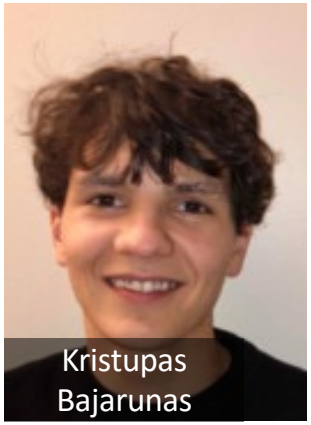
School of Engineering

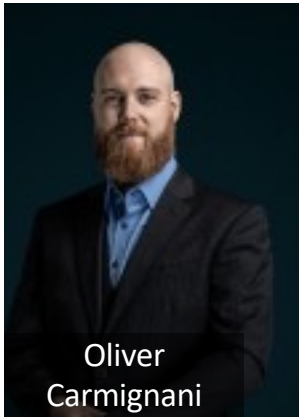
Smart Maintenance Zurich University of Applied Sciences: Engineering + Data science











Wed @13:00 «Physics-informed DL for FD in PV plants»

Project partners

zh aw School of Engineering

Bystronic

swissport

Nispera

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Bundesamt für Strassen ASTRA
Office fédéral des routes OFROU
Ufficio federale delle strade USTRA

STADLER
Clevere Lösungen auf der Schiene

ZÜND
swiss cutting systems

iwb

General Electric

ALSTOM

BÜHLER

EKZ

mespas

am-tec
switzerland

Oxand
Asset, Ageing & Risk Management

SBB CFF FFS

EnBW
Energie braucht Impulse

STADTWERK
WINTERTHUR

NASA

WIN GD
Winterthur Gas & Diesel

MACQUARIE

Coca-Cola

ASE

EH

FLUENCE
A Siemens and AES Company

SMC

Endress+Hauser
People for Process Automation

Operational DL algorithms: a different set of challenges

Nguyen, Medjaher & Tran (2022), A review of artifcial intelligence methods for engineering prognostics and health management with implementation guidelines.
Fink, Wang, Svensen, Dersin, Lee, & Ducoffe (2020), Potential, challenges and future directions for deep learning in prognostics and health management applications.

Selected focus topics

Focus on fault detection and isolation.

I. Few labelled faults.
II. Multi-component systems.
III. Heterogenous fleets of machines.
IV. Scarce training data.
V. Decision under uncertainty.

Generic concepts

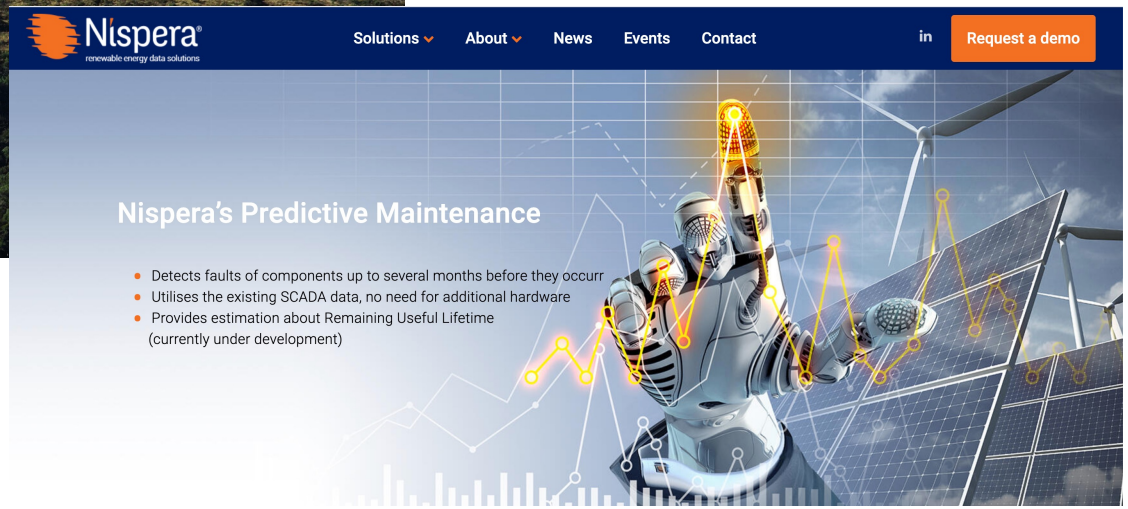
- Regression for anomaly detection and localization.
- Approaches for transfer learning.
- Uncertainty informed anomaly detection.

Example: FDI in wind turbines

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

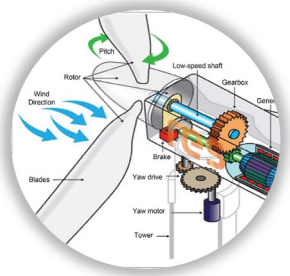
Innosuisse – Swiss Innovation Agency



FLUENCE
A Siemens and AES Company

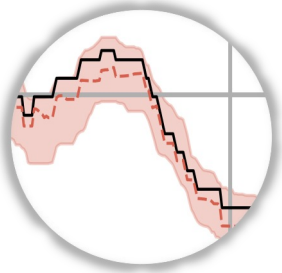
EKZ

More details here:

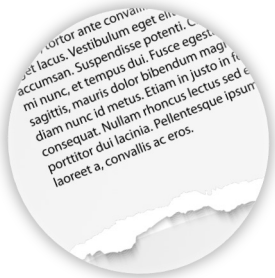


Ulmer, Markus, et al. "Early fault detection based on wind turbine scada data using convolutional neural networks." *5th European Conference of the Prognostics and Health Management Society,*. Vol. 5. No. 1. PHM Society, 2020.

Zgraggen, Jannik, et al. "Transfer Learning Approaches for Wind Turbine Fault Detection using Deep Learning". *6th European Conference of the Prognostics and Health Management Society,* Vol. 6. No. 1. PHM Society, 2021. ******Best Paper Award***

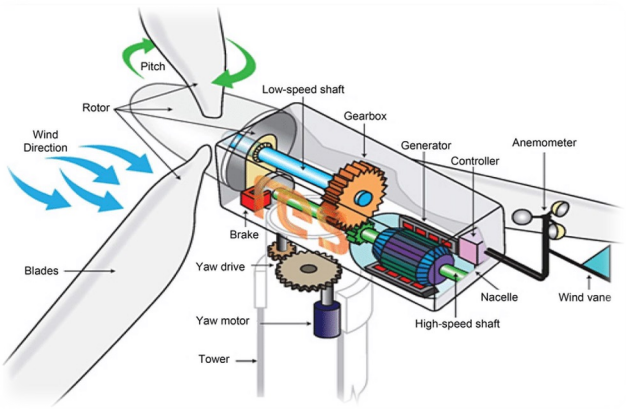


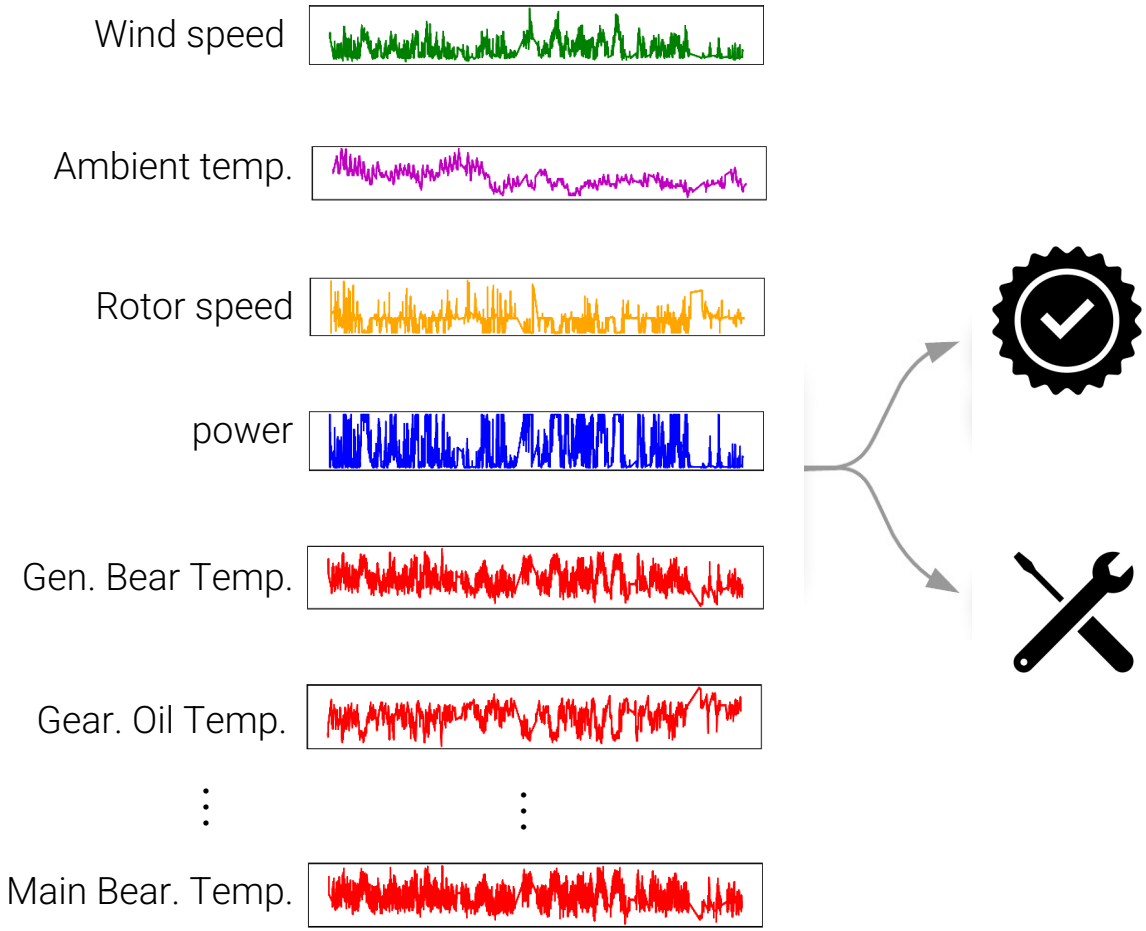
Zgraggen, Jannik, et al. "Uncertainty informed anomaly scores with deep learning: robust fault detection with limited data". *7th European Conference of the Prognostics and Health Management Society,* Vol. 7. No. 1. PHM Society, 2022. ******Best Paper Award***



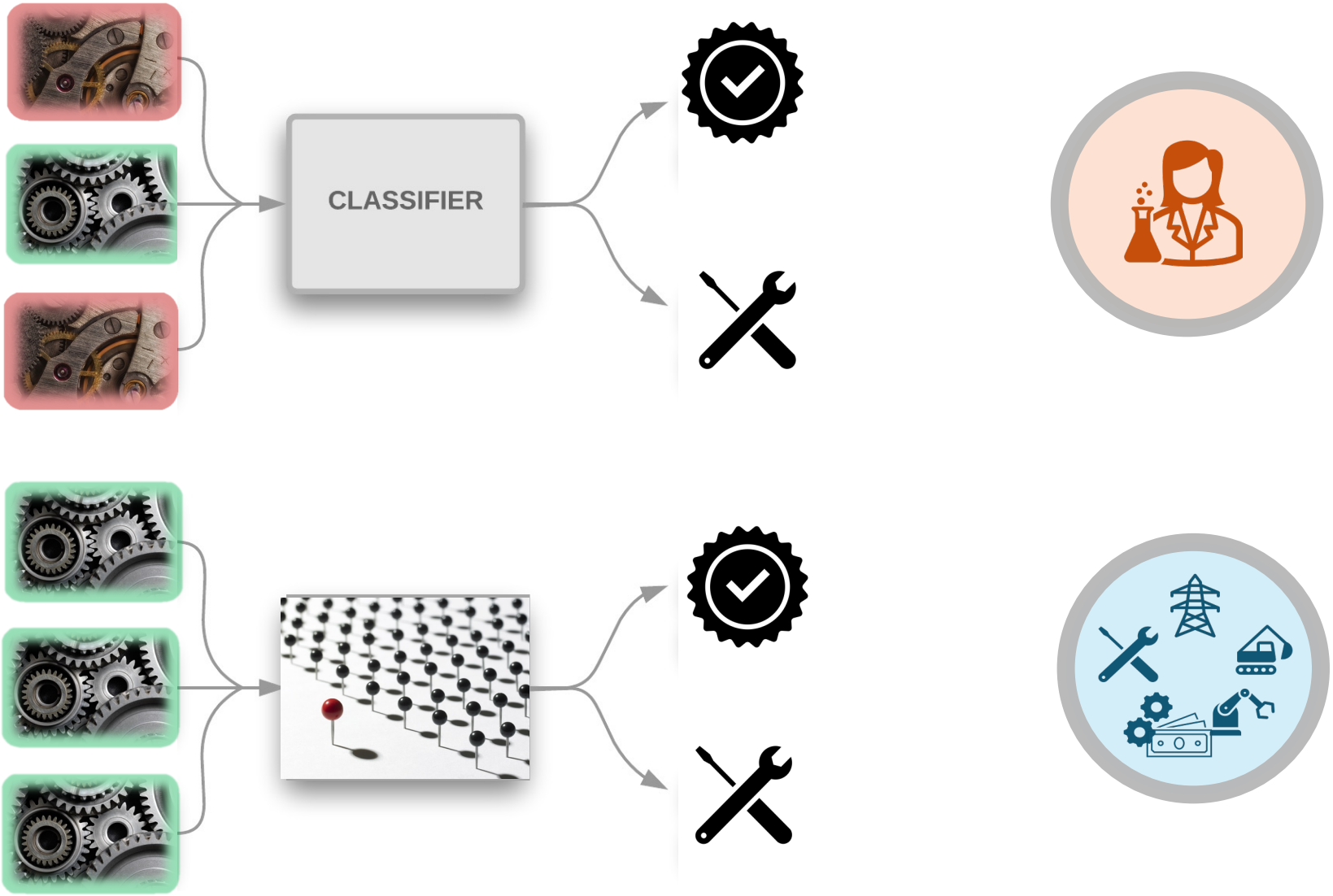
Ulmer, Markus, et al. "Scaling up Deep Learning Based Predictive Maintenance for Commercial Machine Fleets: a Use Case", Proceedings of the SDS Conference 2022. ******Best Presentation Award**

Goal: detect faults with existing SCADA data

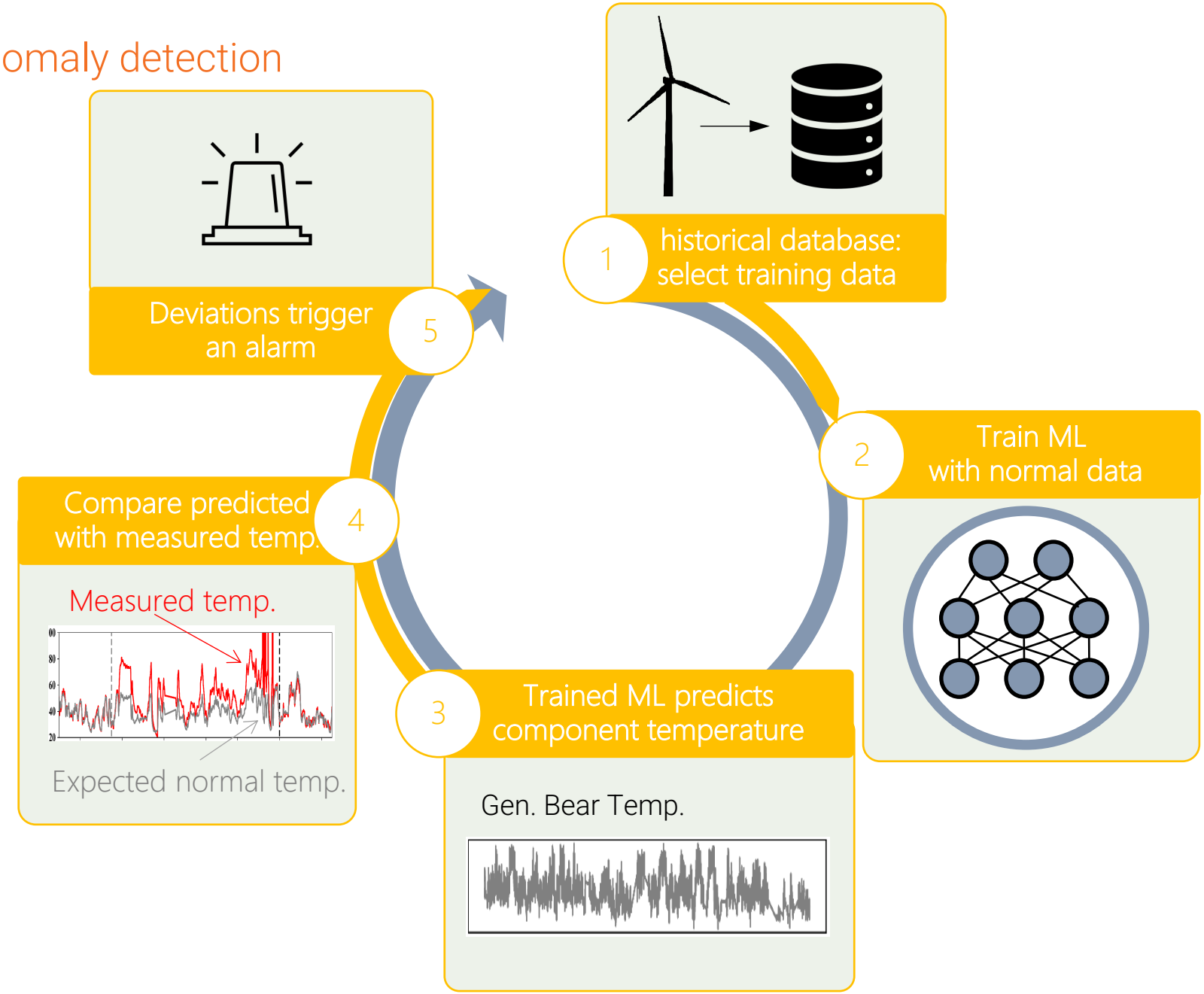




Gap I: fault detection without labeled faults

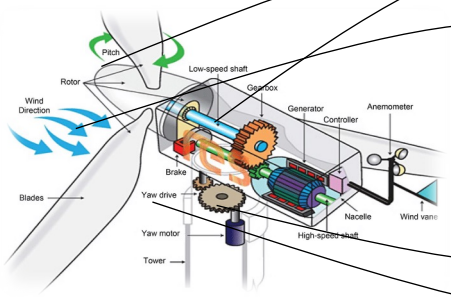


Anomaly detection



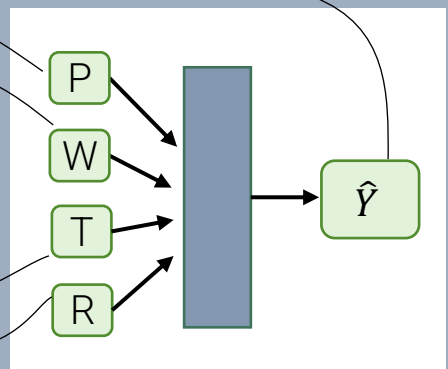
Regression for anomaly detection

Data



Scada data (10min)

Regression model



Healthy state model

Anomaly Detection

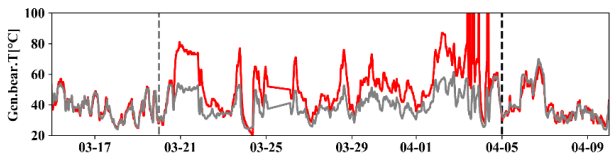
$$r = Y - \hat{Y}$$

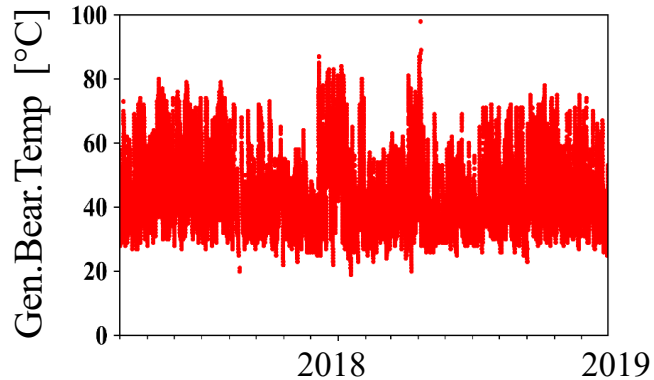
Compute anomaly score for each data point

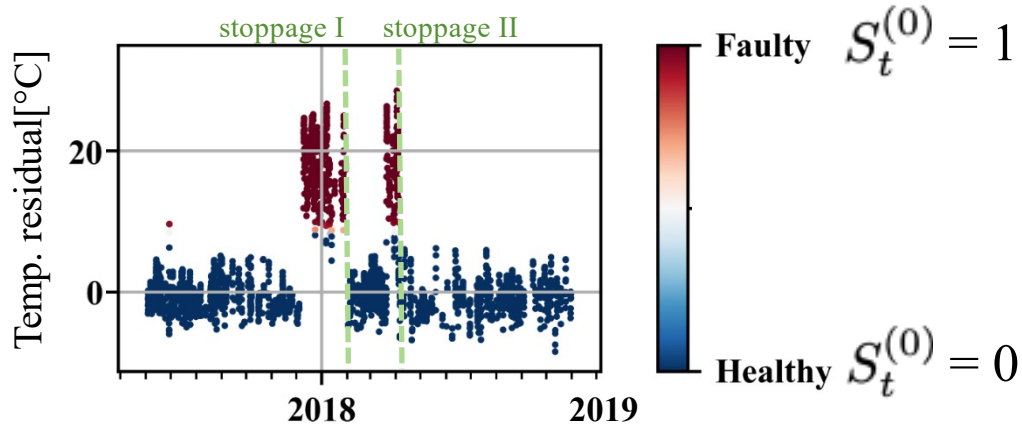
Fault Detection: Application and Validation

Example:

Generator bearing temp. in a 2MW turbine



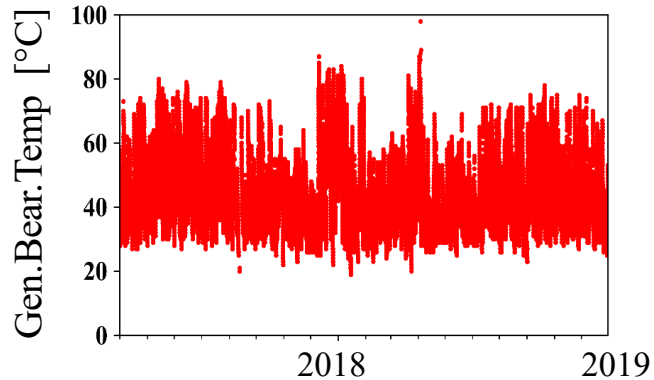


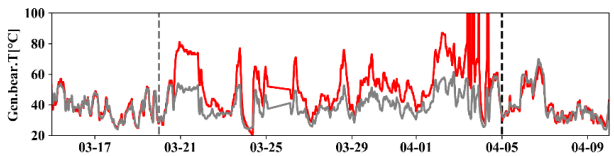


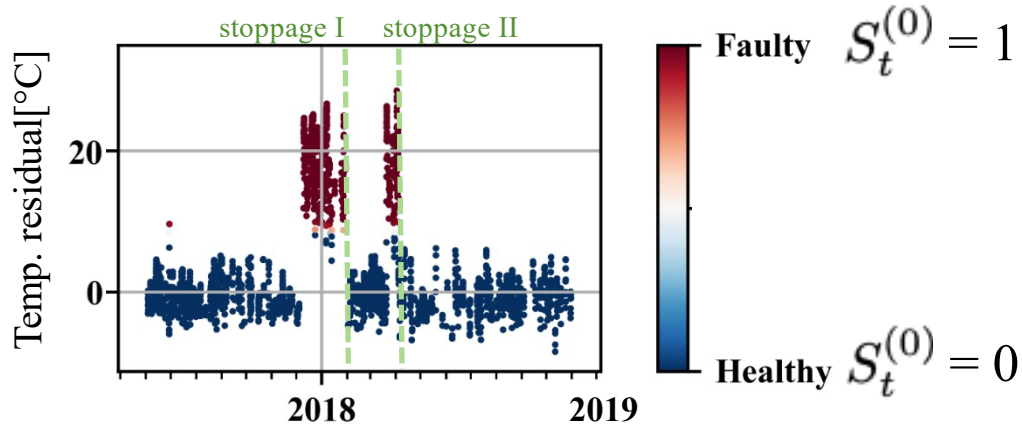
Fault Detection: Application and Validation

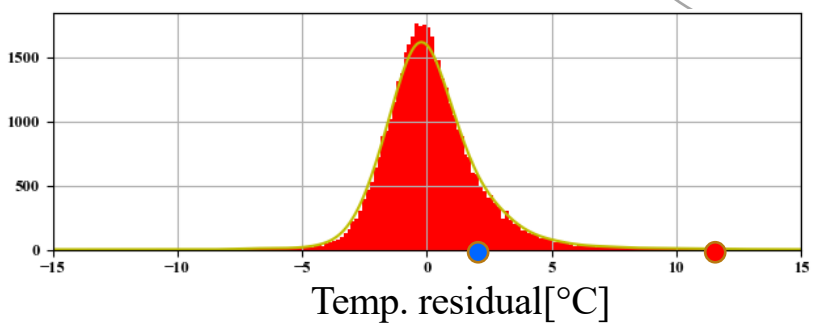
Example:

Generator bearing temp. in a 2MW turbine







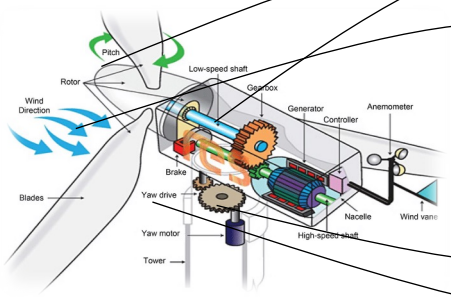


$$S_t^{(0)} = F(r_t; \mu^{(tr)}, \sigma^{(tr)})$$

Anomaly Score

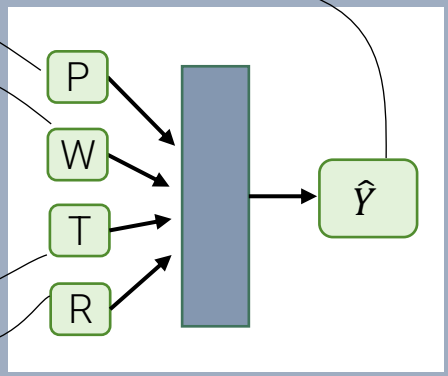
Regression for anomaly detection: recap

Data

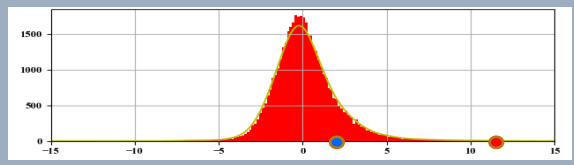


Scada data (10min)

Regression model



Healthy state model



Temp. residual[°C]

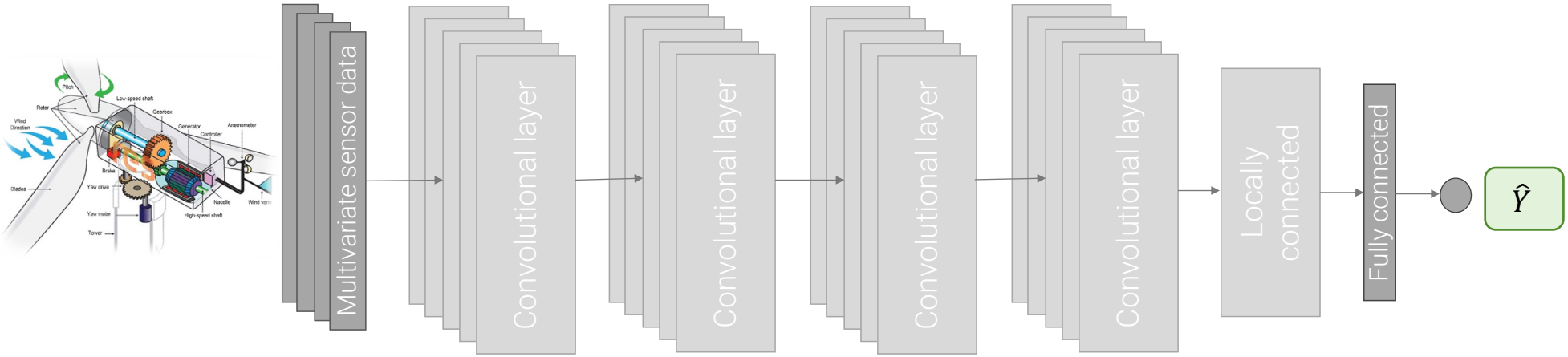
Anomaly Detection

$r = Y - \hat{Y}$

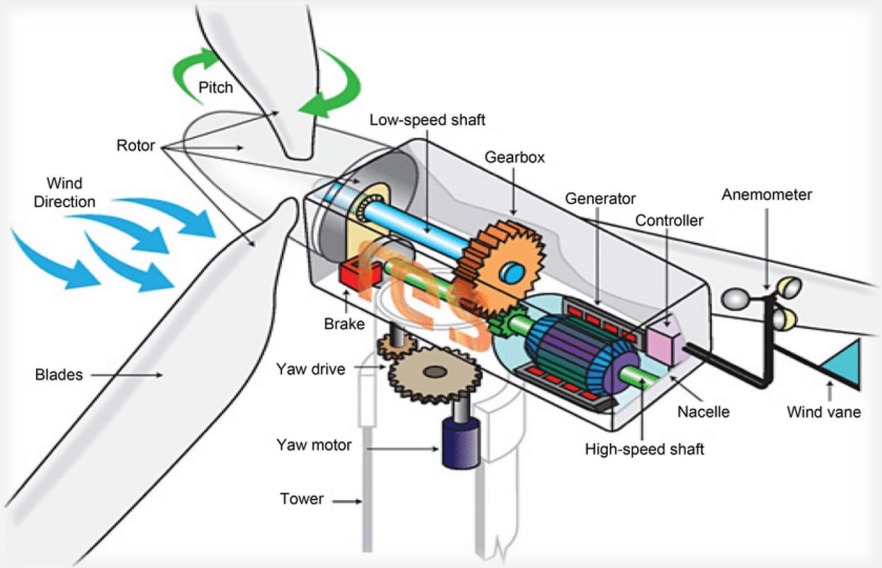
Compute anomaly score for each data point

$$S_t^{(0)} = F(r_t; \mu^{(tr)}, \sigma^{(tr)})$$

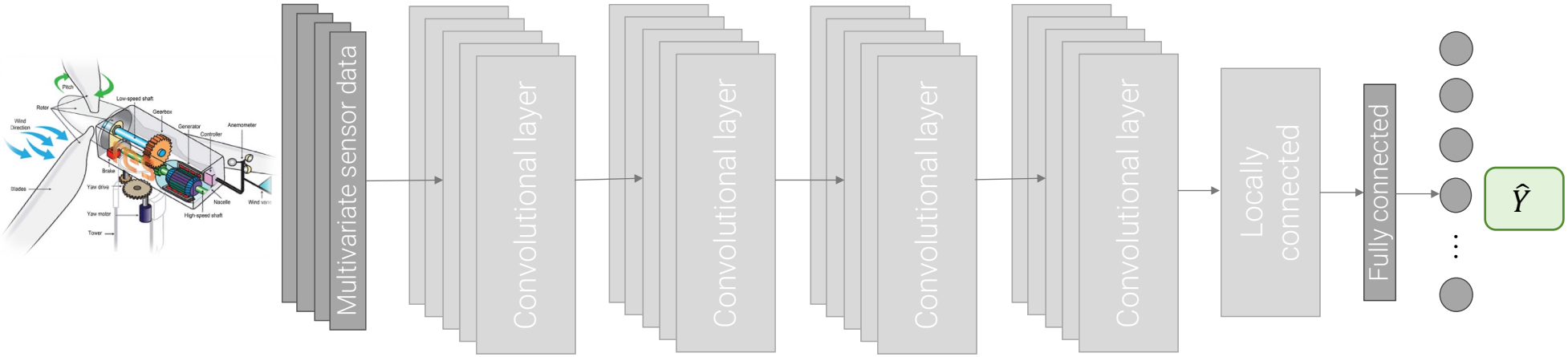
Regression with CNN



Gap II: scaling to multi-component systems

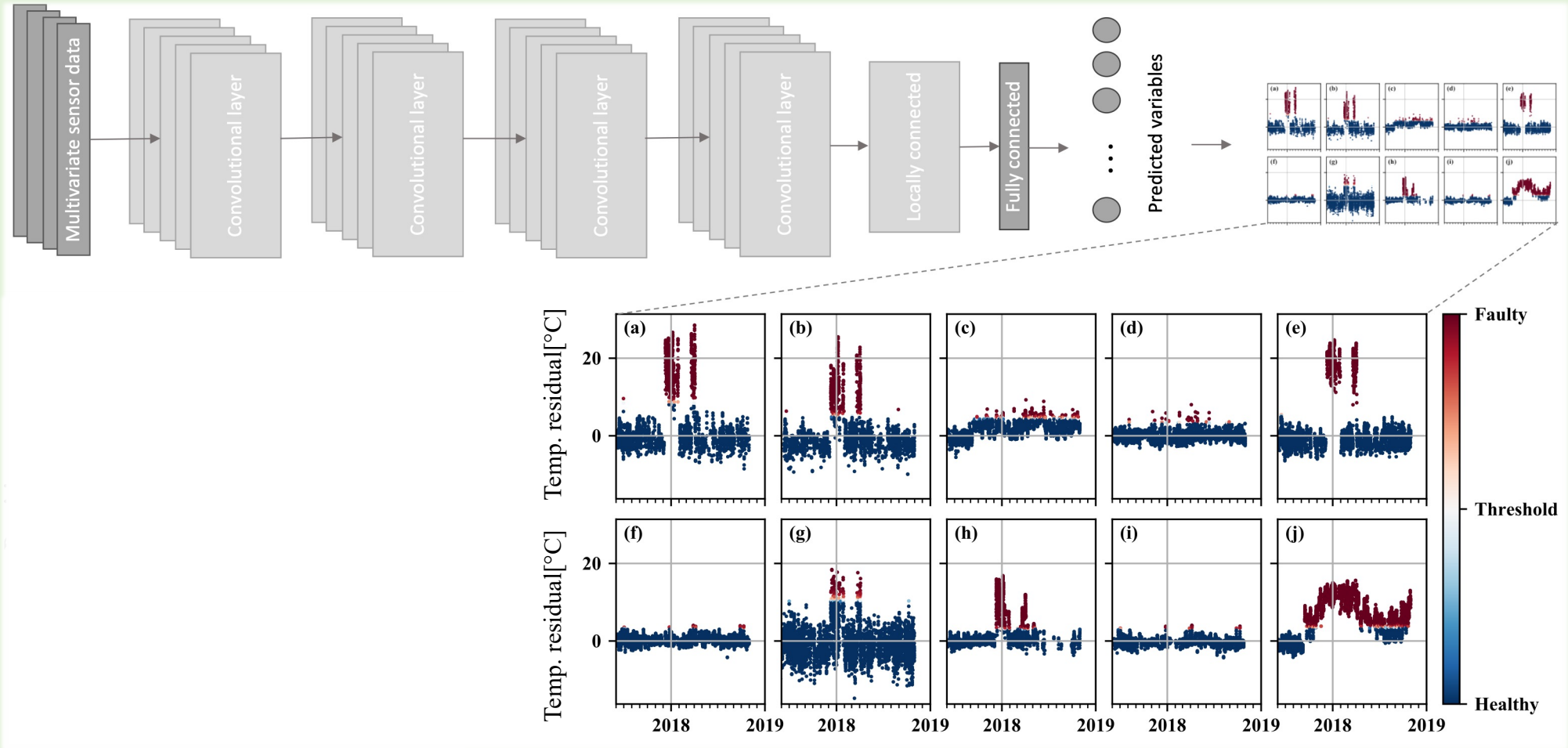


Multi-output regression

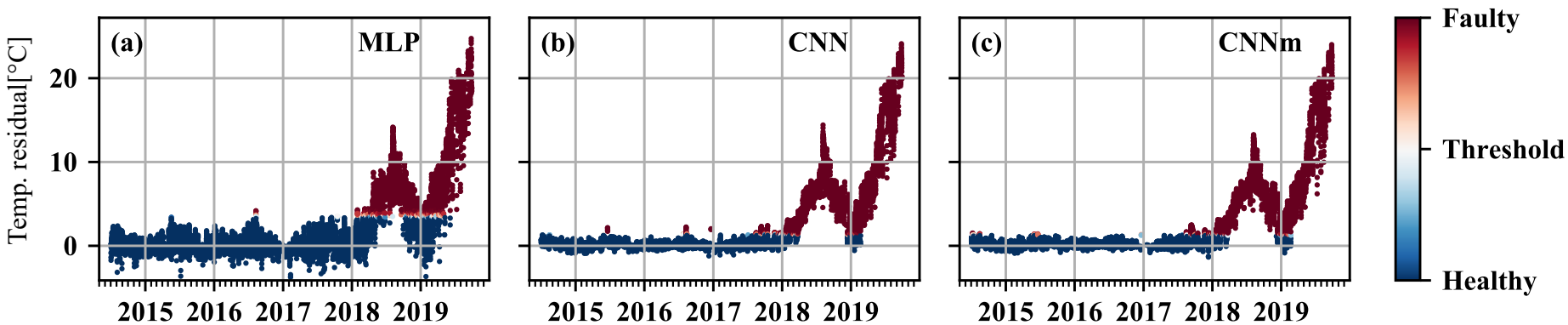


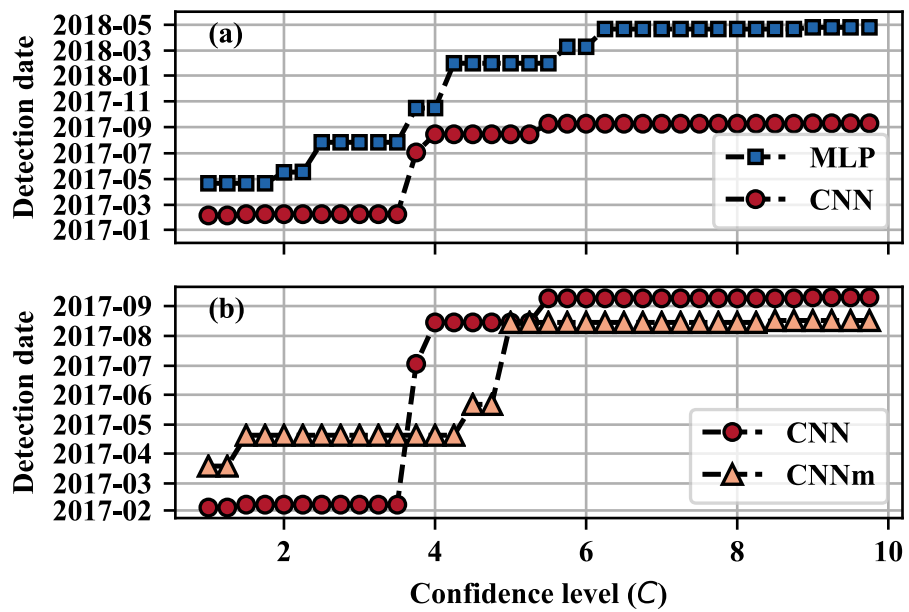
Ulmer, Markus, et al. "Early fault detection based on wind turbine scada data using convolutional neural networks." PHM Society, 2020.

Multi-output regression



Validation: slow degradation



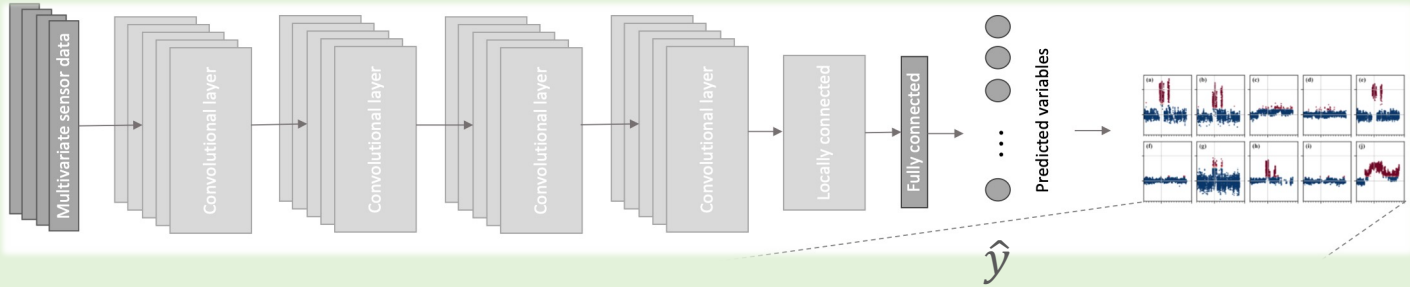


Ulmer, Markus, et al. "Early fault detection based on wind turbine scada data using convolutional neural networks." PHM Society, 2020.

The advantage of regression for fault isolation

regression

X



$\hat{y}$

$$r_j = y_j - \hat{y}_j$$

reconstruction

Wind speed

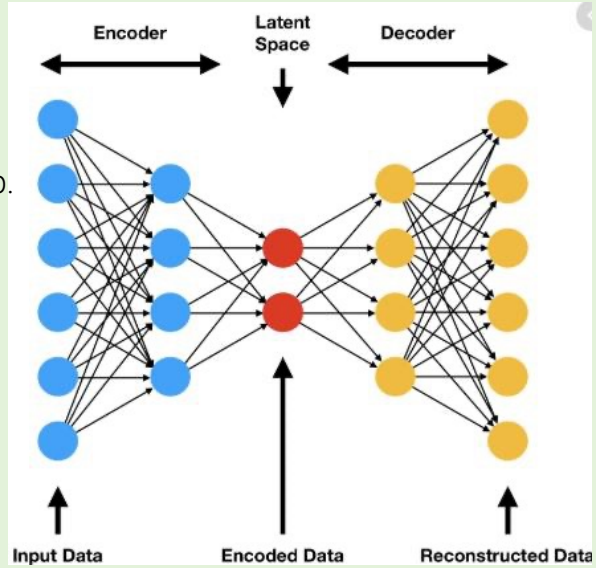
Ambient temp.

Rotor speed

power

Gen. Bear Temp.

Gear. Oil Temp.



X $\hat{X}$

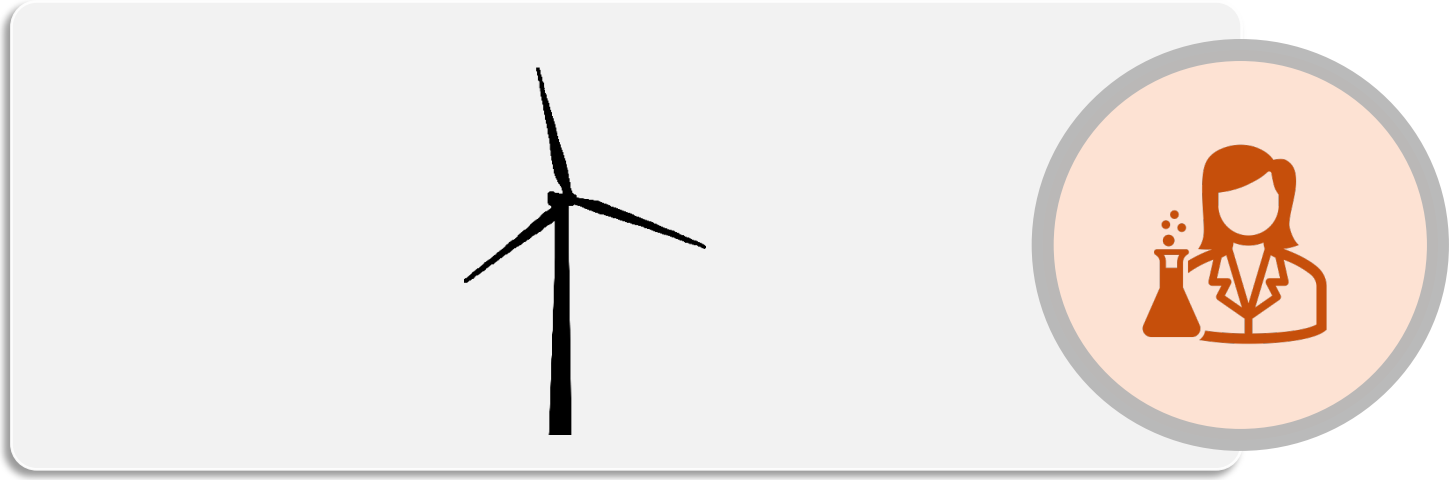
$$r_j = X_j - \hat{X}_j$$

Disadvantage:

Error spill-over ➔ no fault localization

Z Yang, et al (2021)."Autoencoder-based representation learning and its application in intelligent fault diagnosis: A review

Gap III: scaling to heterogenous machine fleets

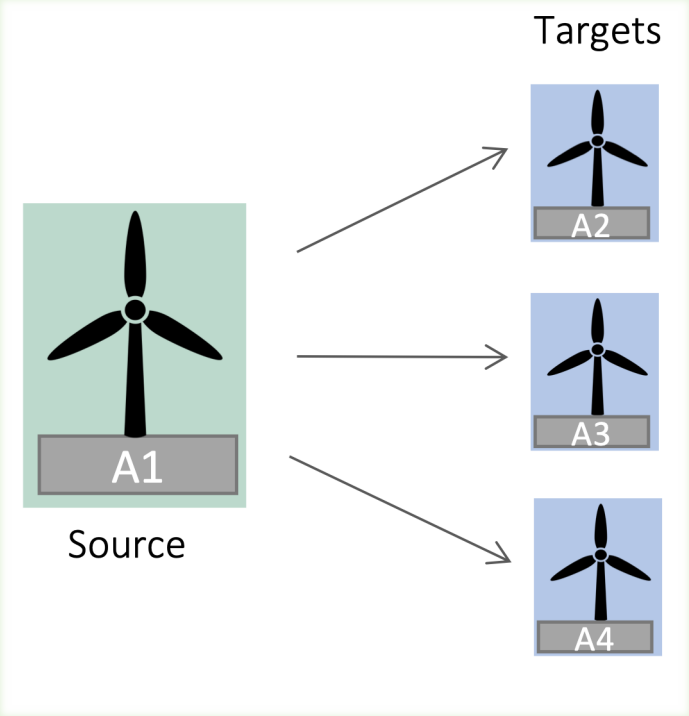


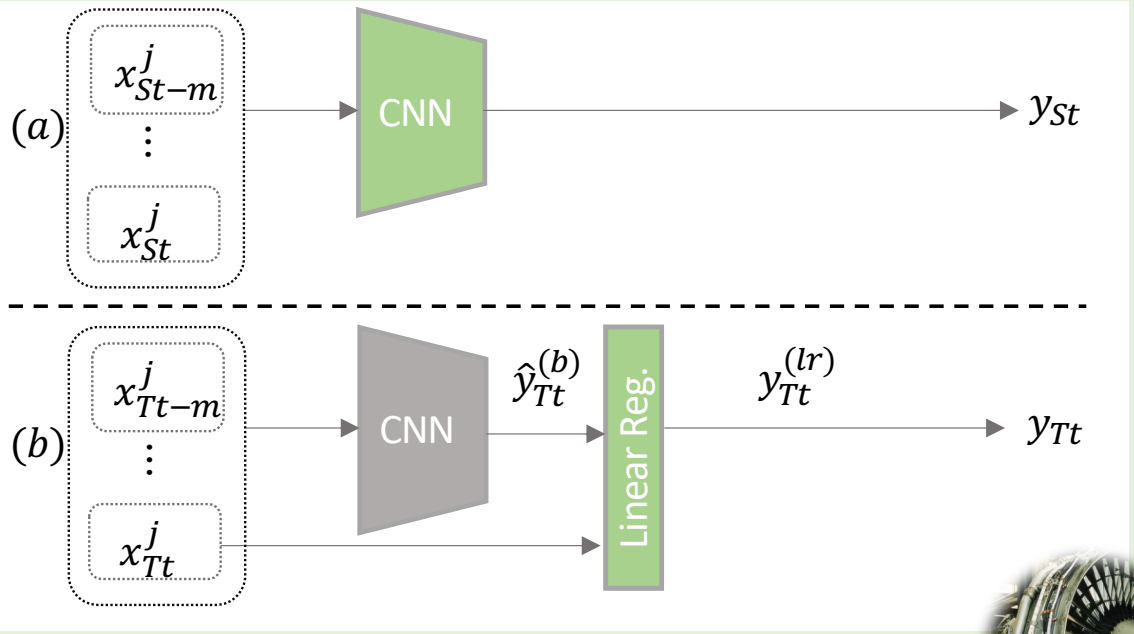
Gap III: scaling to heterogenous machine fleets

selecting the right method is application-specific!

Focus: “real world” evaluation approaches.

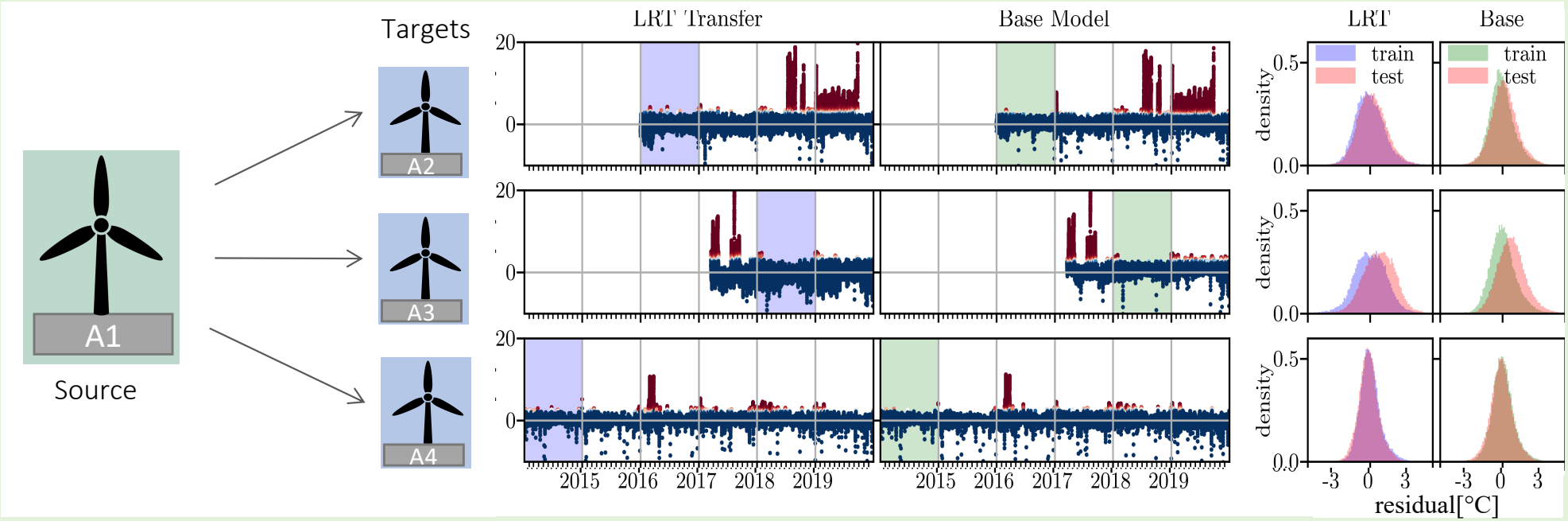
Tranfer learning for scaled training





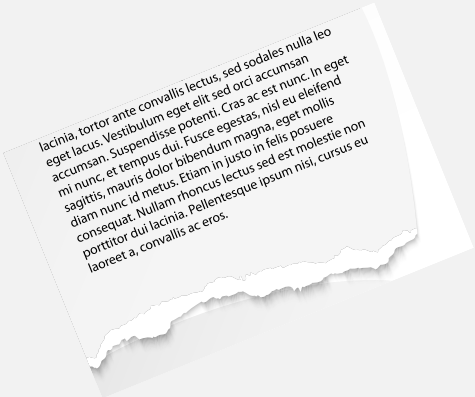
Domain shift in target variable (unit-to-unit).

Tranfer learning for scaled training: how to evaluate?

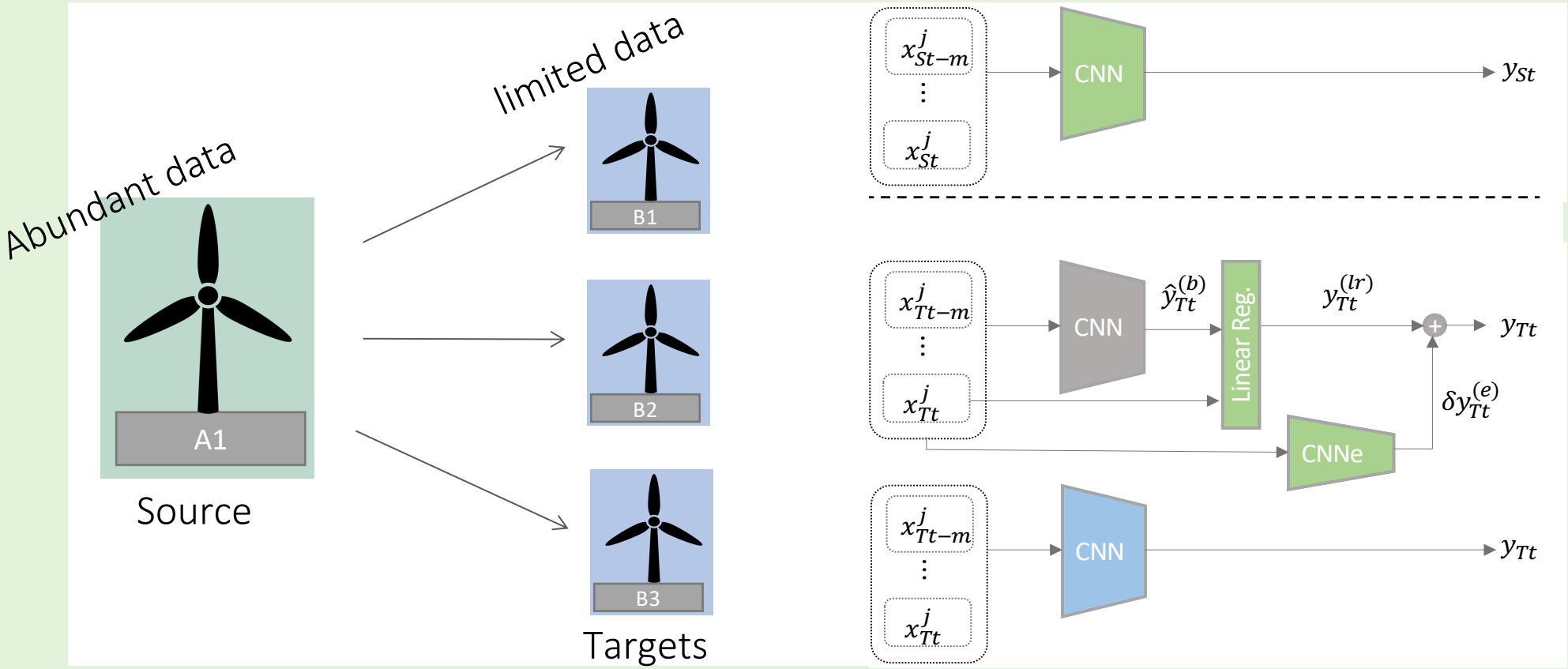


Compare the residual distribution shift between train and test.

Gap IV: solution for limited data scenario



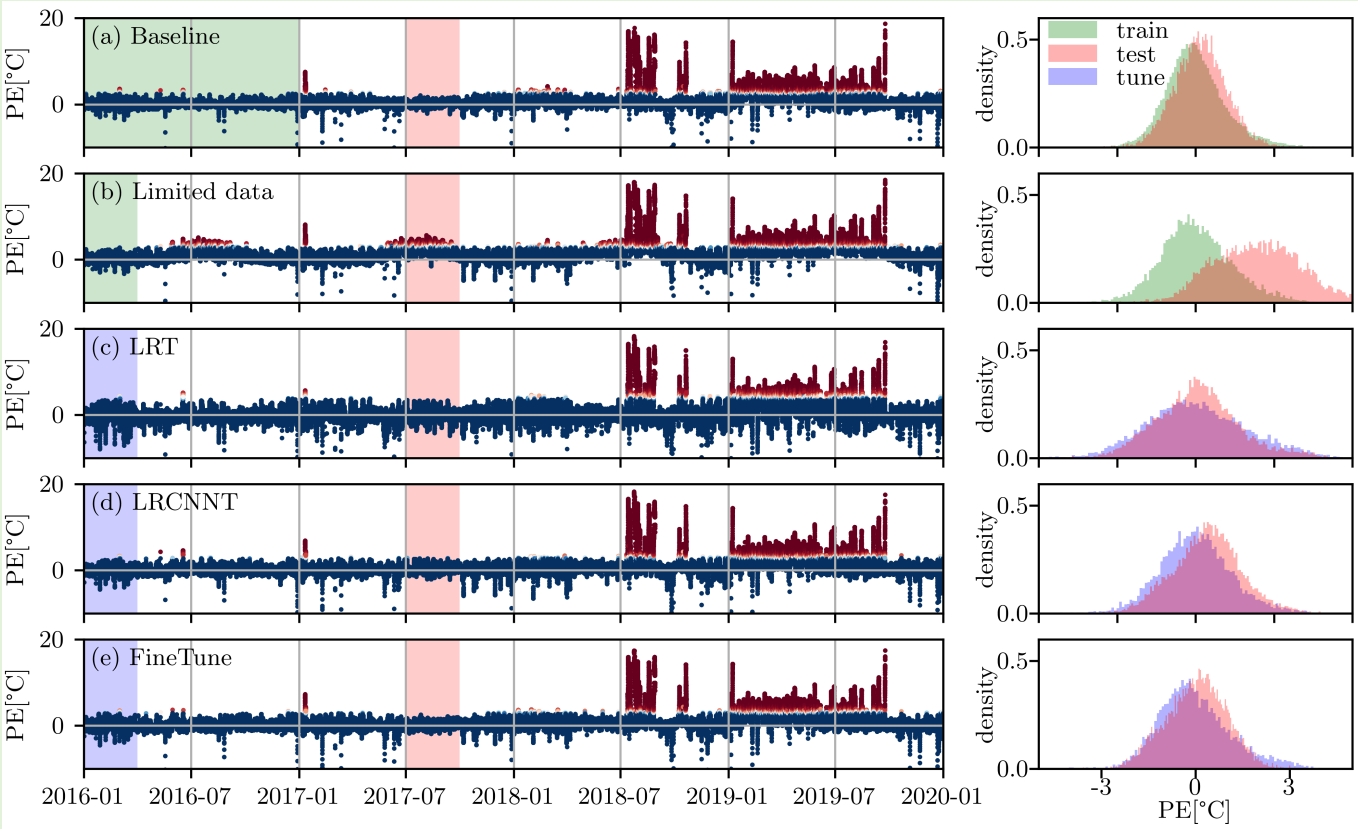
Transfer Learning for Data Scarcity



Domain shift in (i) target variable (unit-to-unit).
(ii) inputs (fleet-to-fleet).
(iii) operative conditions (winter to summer).

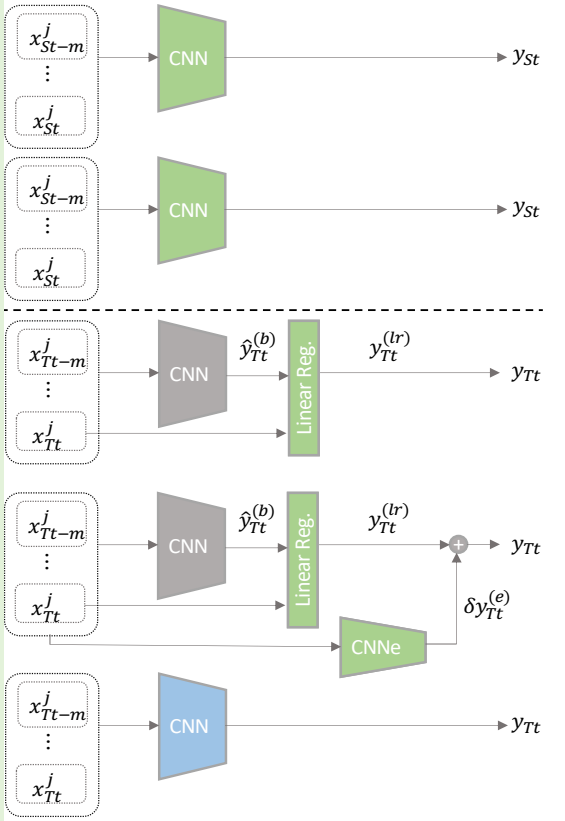
Transfer Learning for Data Scarcity: how to evaluate?

Result



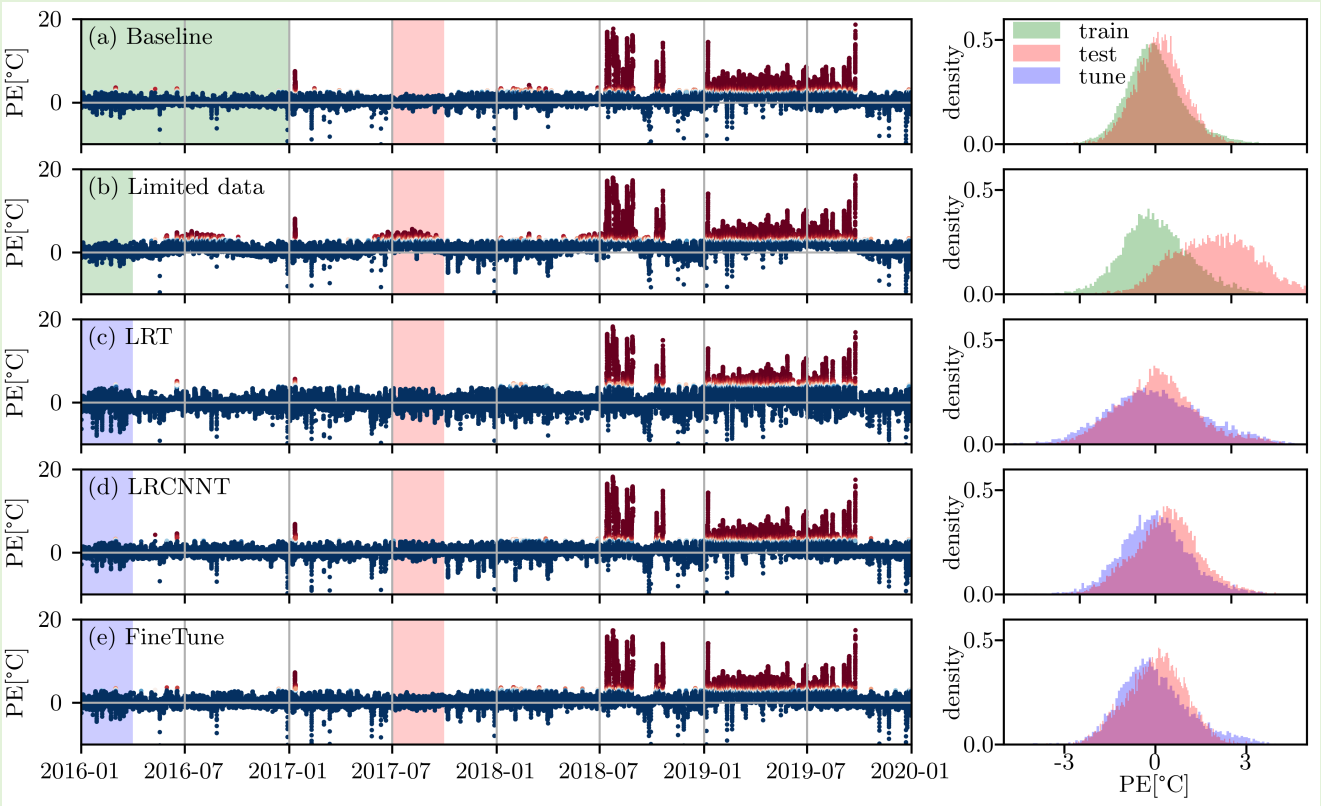
Evaluation

Method



Compare the residual distribution shift between train and test.

Transfer Learning for Data Scarcity: how to evaluate?

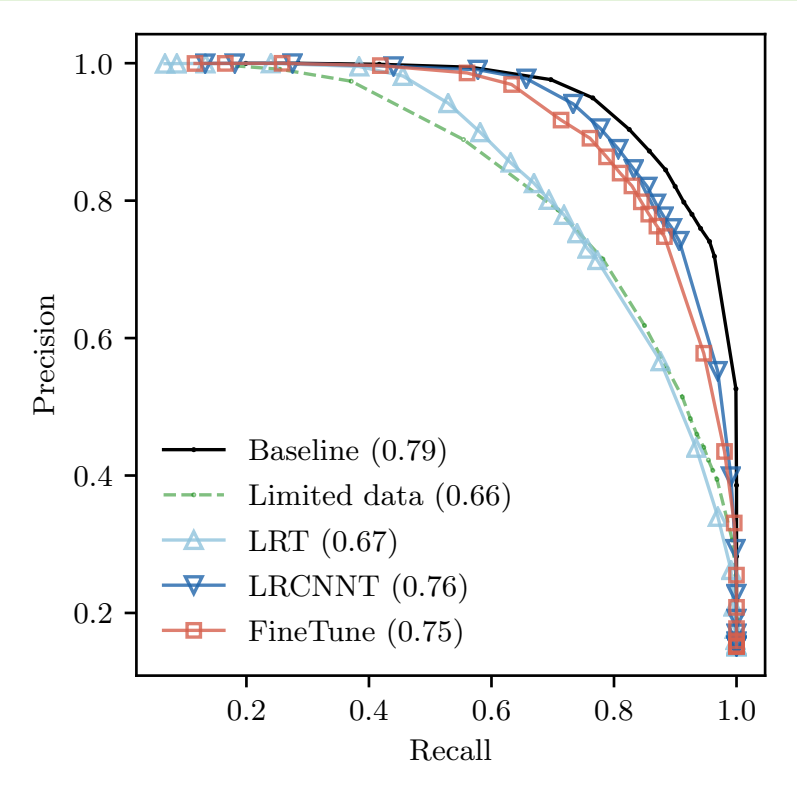


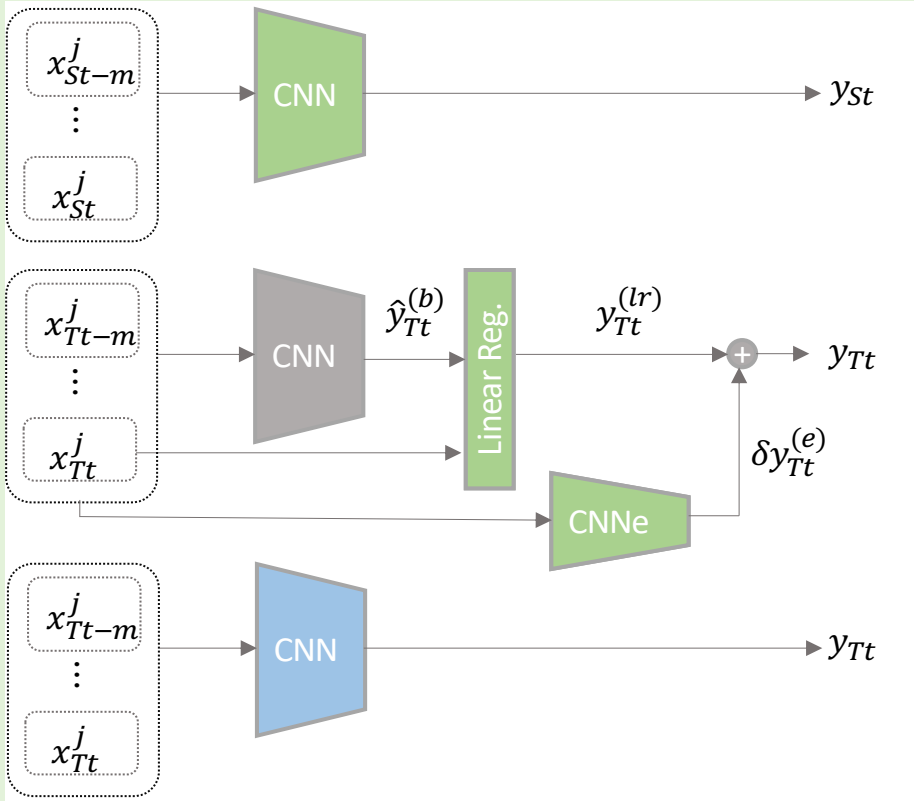
Model	μ shift	σ shift
Base Model	0.2 ± 0.03	-0.1 ± 0.02
Limited Data	1.46 ± 0.38	0.65 ± 0.32
LRT	-0.1 ± 0.11	-0.34 ± 0.05
LRCNNT	0.22 ± 0.09	-0.09 ± 0.04
FineTune	-0.24 ± 0.18	-0.15 ± 0.07

Zgraggen Jannik, et al. " Transfer Learning Approaches for Wind Turbine Fault Detection using Deep Learning" PHME Society, 2021.
BEST PAPER AWARD

Compare the residual distribution shift between train and test.

Transfer Learning for Data Scarcity: how to evaluate?

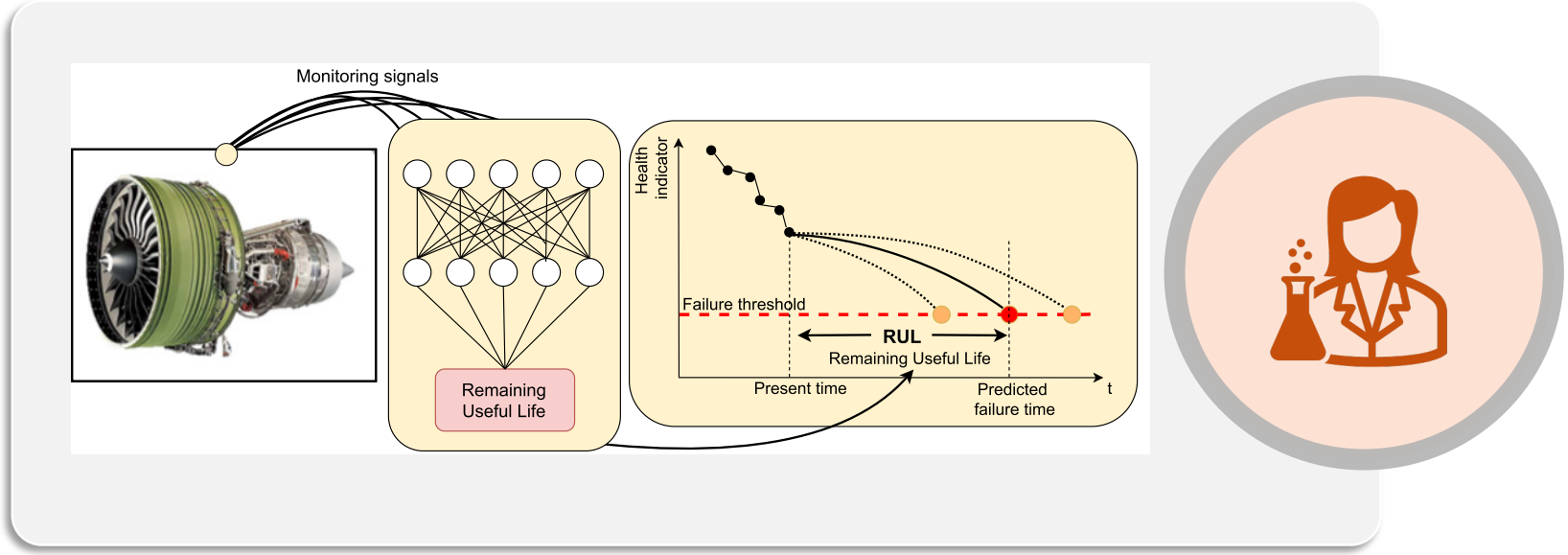


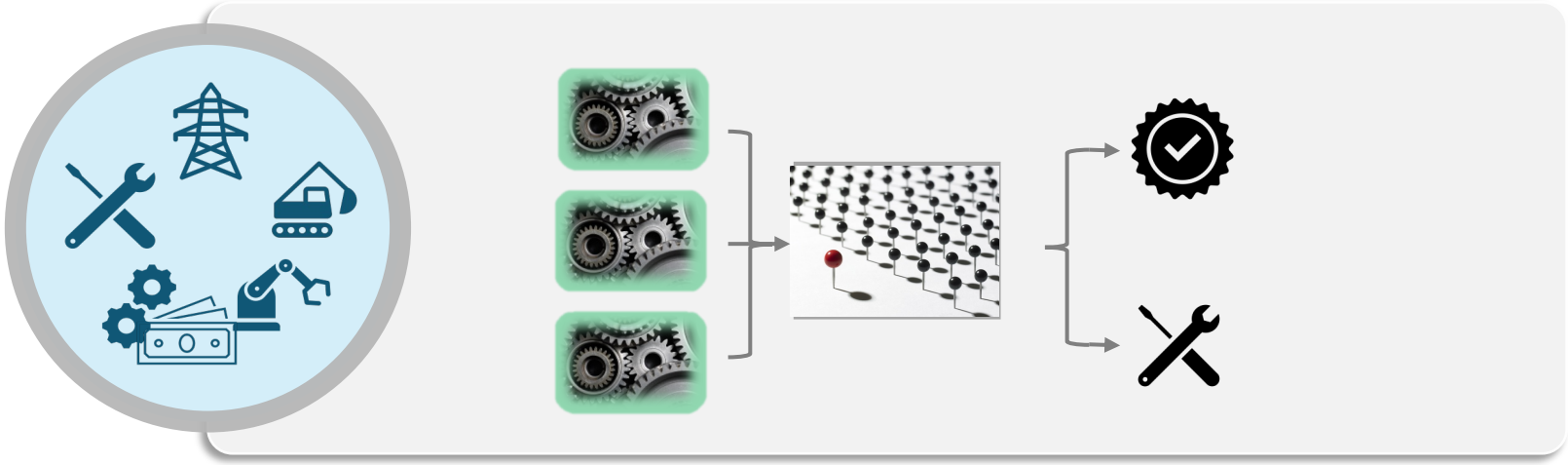


Use a model you trust as a baseline for labeling.

Zgraggen Jannik, et al. " Transfer Learning Approaches for Wind Turbine Fault Detection using Deep Learning" PHME Society, 2021. BEST PAPER AWARD

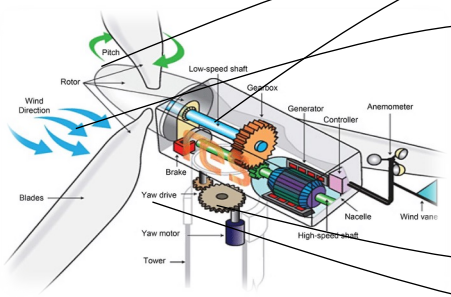
Gap V: uncertainty quantification (also for anomaly detection)





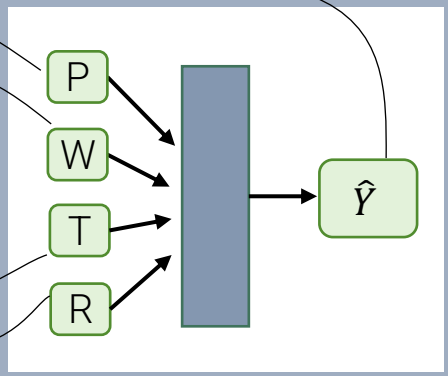
Regression for anomaly detection: recap

Data

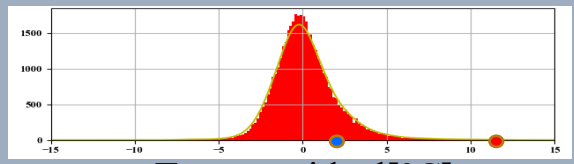


Scada data (10min)

Regression model



Healthy state model



Temp. residual[°C]

Anomaly Detection

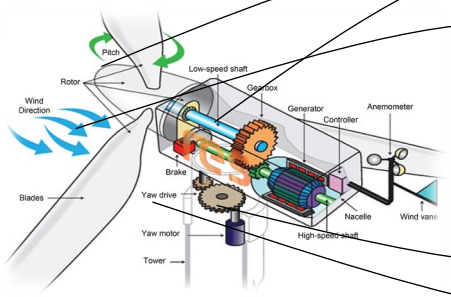
$r = Y - \hat{Y}$

Compute anomaly score for each data point

$$S_t^{(0)} = F(r_t; \mu^{(tr)}, \sigma^{(tr)})$$

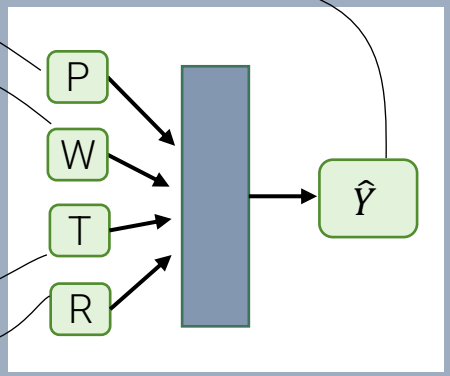
Regression for anomaly detection: recap

Data



Scada data (10min)

Regression model



Healthy state model

Quantify Uncertainty

Anomaly Detection

$r = Y - \hat{Y}$

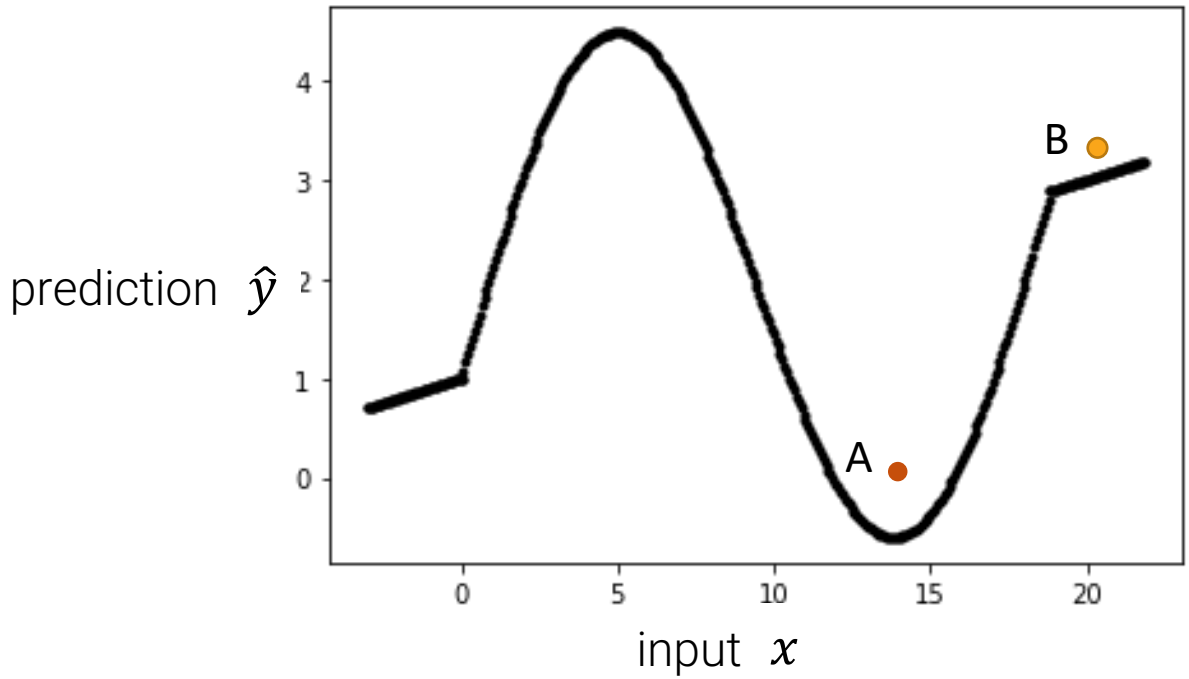
Compute anomaly score for each data point

Use Uncertainty

Step 1: Uncertainty quantification
Normal state models for anomaly detection

i. Model based on training data.

ii. Decide if new measurements are anomalies.



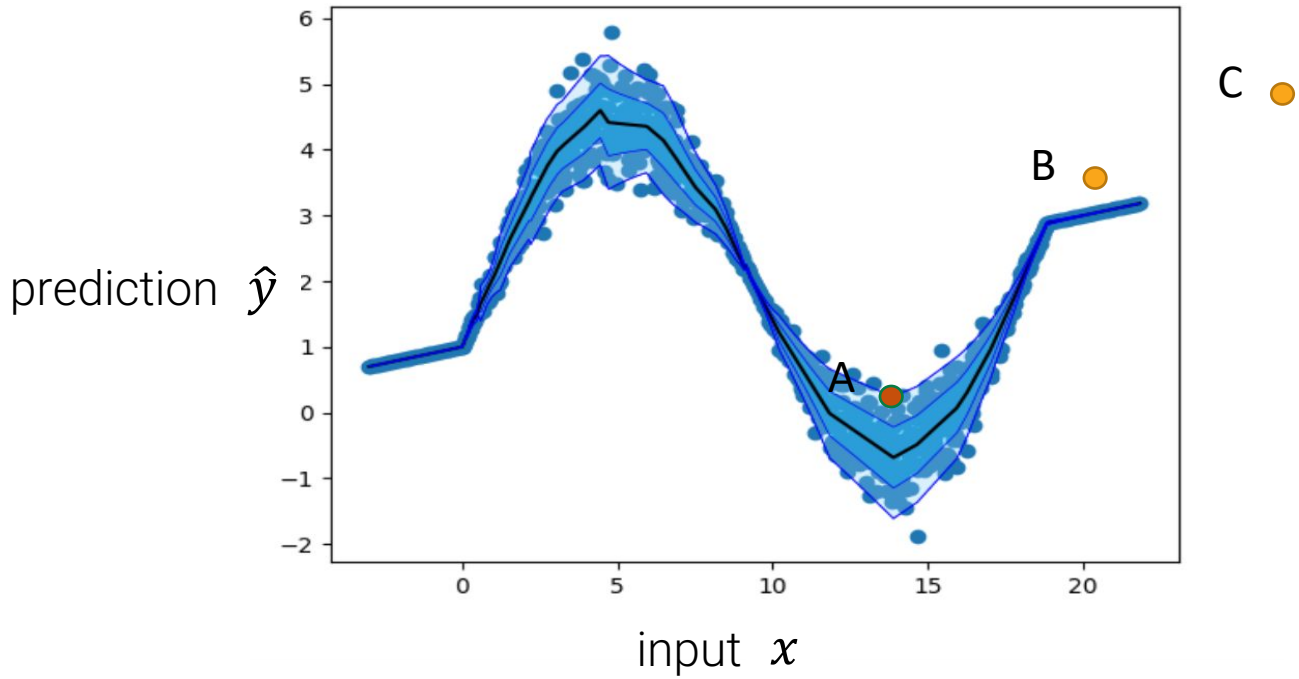
Step 1: Uncertainty quantification
Normal state models for anomaly detection

i. Model based on training data.

ii. Decide if test data A and B are anomalies.

➔Can we quantify the uncertainty of our predictions?

- Aleatoric uncertainty
- Epistemic uncertainty

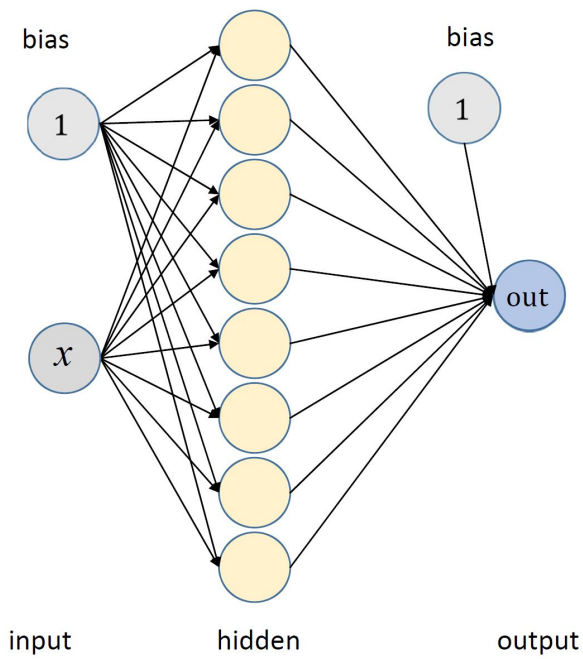


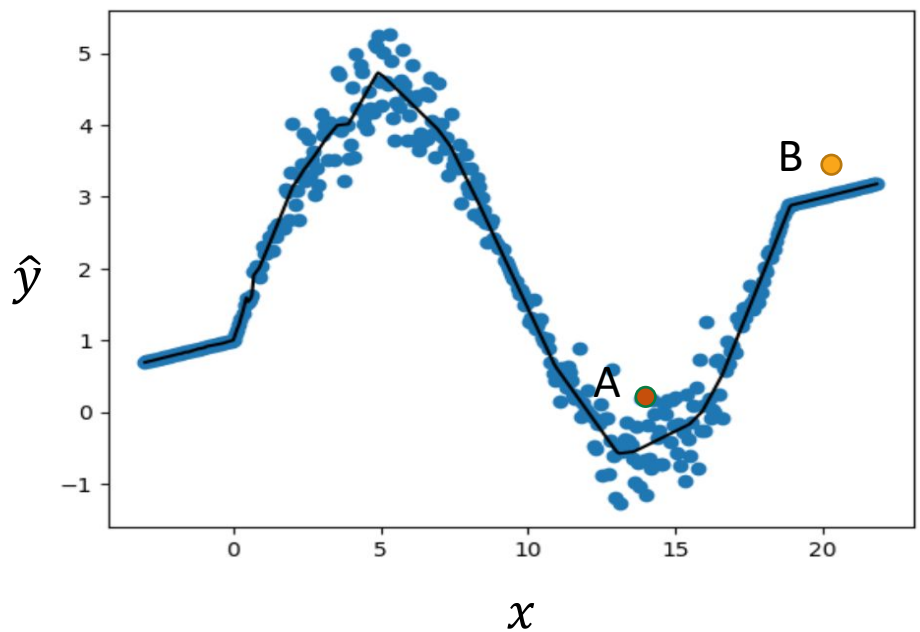
Aleatoric uncertainty (data uncertainty)

Standard regression with MSE loss

Implicit assumption: heteroscedasticity

$$L_{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$



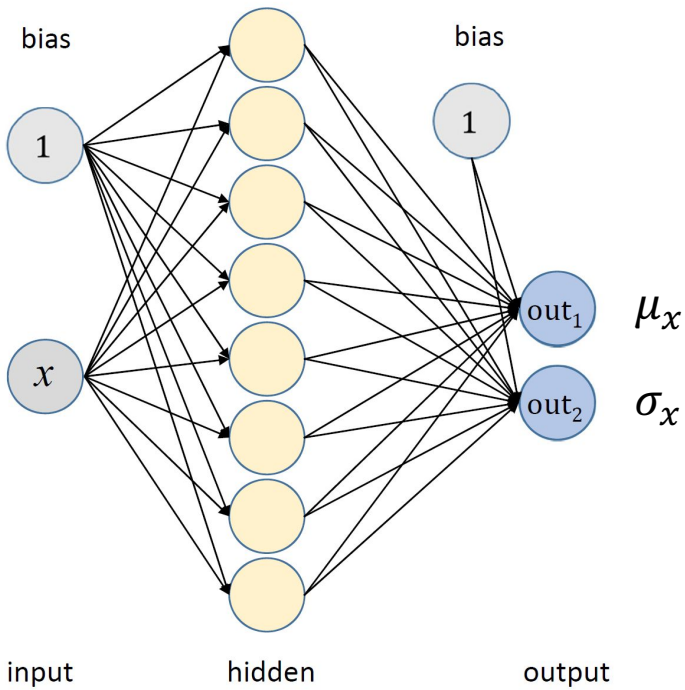


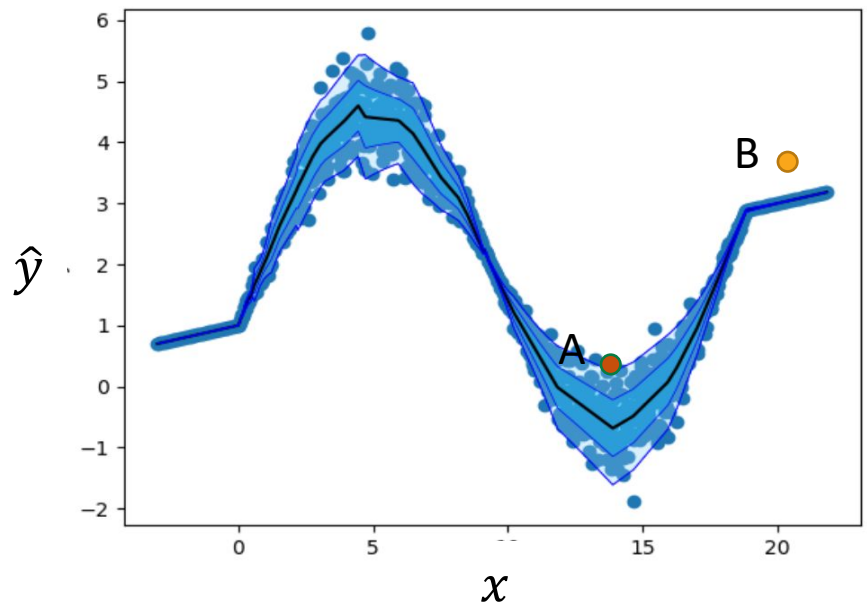
How to model aleatoric uncertainty in a NN?

Distribution modelling with Negative Log Likelihood loss

➔Input dependent variance = a distribution of predictions

$$L_{NLL} = \sum_{i=1}^{n} -\log\left(\frac{1}{\sqrt{2\pi\sigma_{x_i}^{2}}}\right) + \frac{(\mu_{x_i} - y_i)^2}{2\sigma_{x_i}^{2}}$$





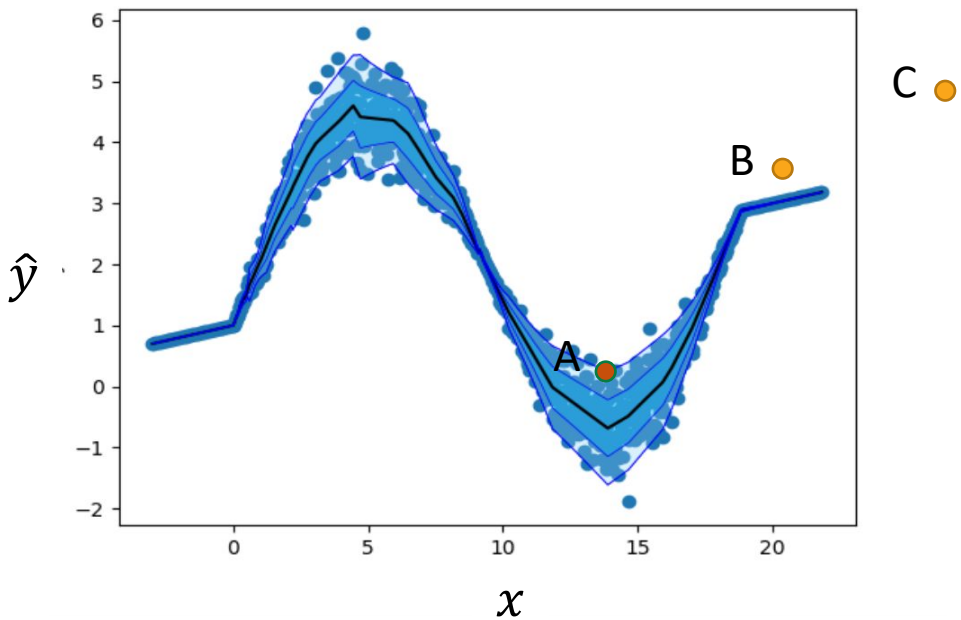
Dürr, O., Sick, B., & Murina, E. (2020). *Probabilistic deep learning: With python, keras and tensorflow probability*. Manning Publications

Epistemic uncertainty (model uncertainty)

Point C is Out of Distribution (OoD) of the training data.

Quantify the model limitation «epistemic uncertainty»

➔ Various methods.



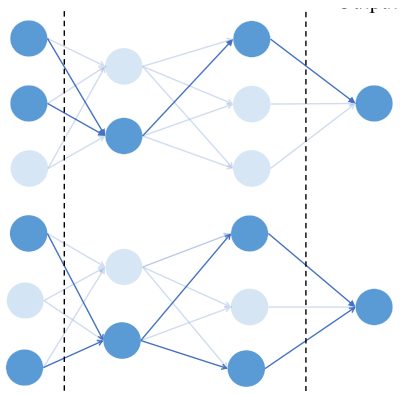
NN-methods for quantifying model uncertainty

Goal:

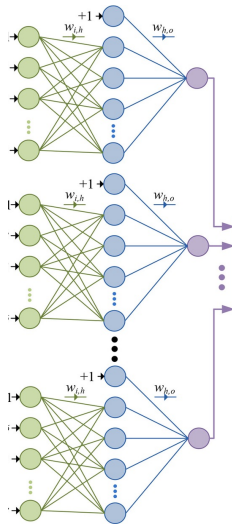
- train a set of models.
- Each suggests a possible solution to the regression problem.
- How different are the predictions for a given input?

➔ $\tilde{\mu}_t$ and $\tilde{\sigma}_t^2$

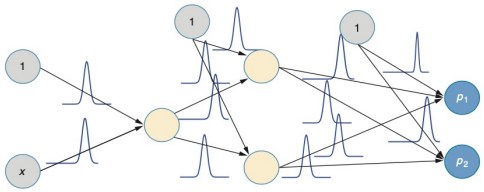
MC-Dropout



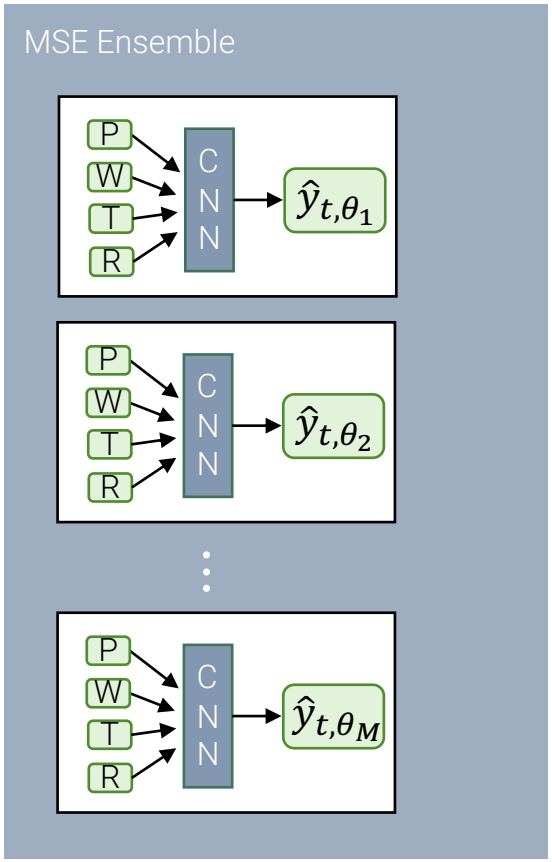
Deep Ensembles

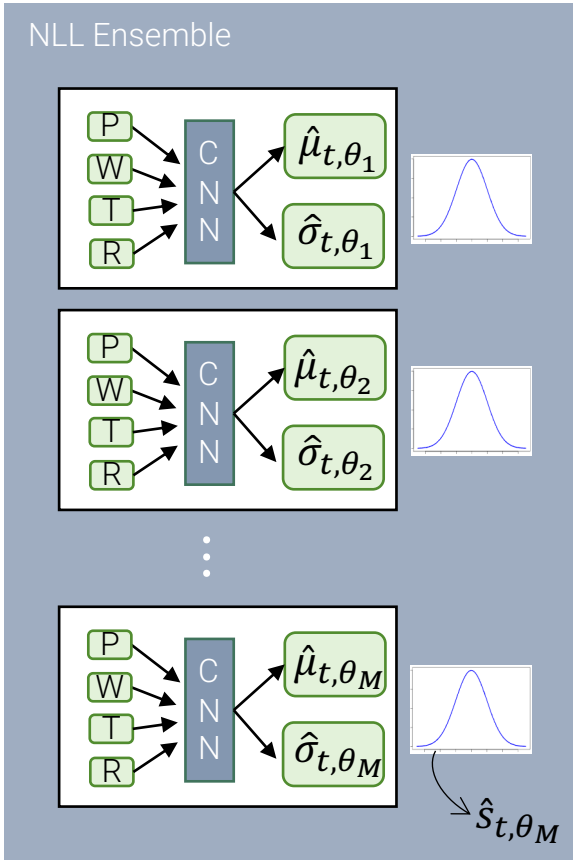


Variational Inference



Deep Ensembles





$$\tilde{\mu}_t = \frac{1}{M}\sum_{m=1}^{M} \hat{y}_{t,\theta_m}$$

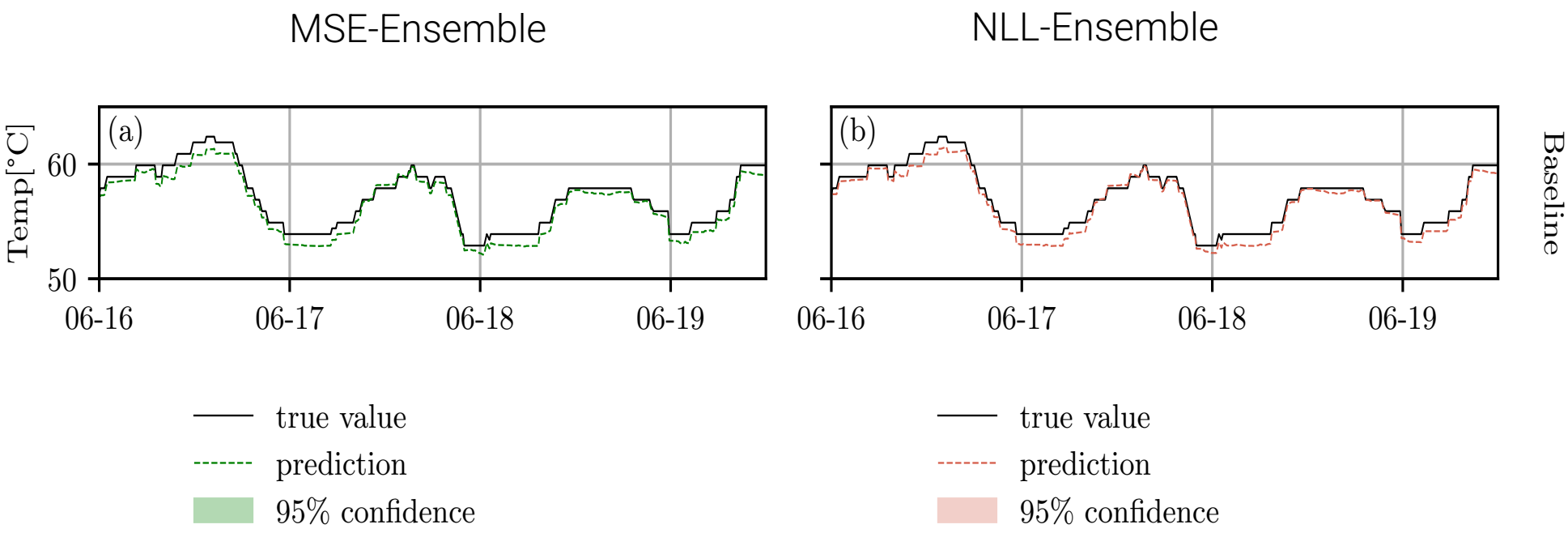
$$\tilde{\sigma}_t^2 = \frac{1}{M-1}\sum_{m=1}^{M} (\hat{y}_{t,\theta_m} - \tilde{\mu}_t)^2$$

$$\tilde{\mu}_t = \frac{1}{M}\sum_{m=1}^{M} \hat{s}_{t,\theta_m}$$

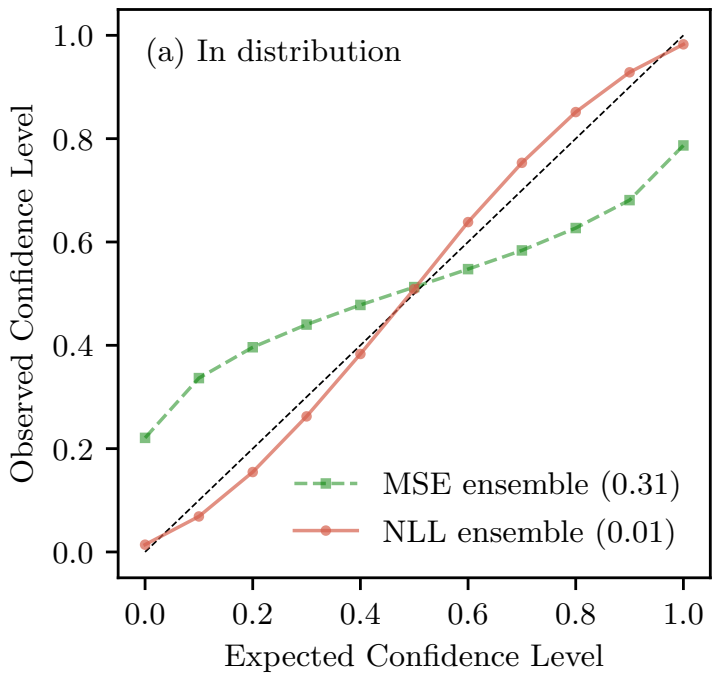
$$\tilde{\sigma}_t^2 = \frac{1}{M-1}\sum_{m=1}^{M} (\hat{s}_{t,\theta_m} - \tilde{\mu}_t)^2$$

Lakshminarayanan, et al. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles.

Results on healthy data: prediction intervals



Calibration curve: selecting an appropriate uncertainty quantification



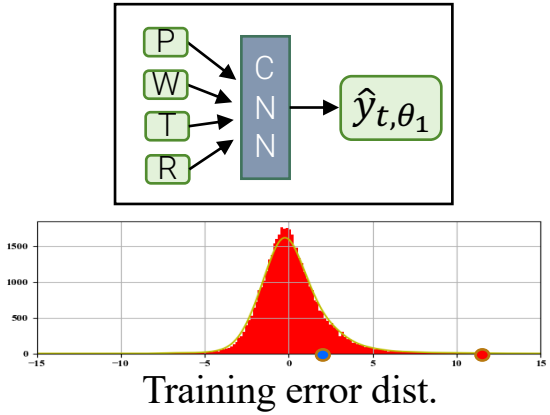
Calibration error:

$$\epsilon_{\text{cal}} = \sum_{j=1}^{n} (p_j - \hat{p}_j)^2.$$

Kuleshov et al., accurate uncertainties for deep learning using calibrated regression. (2018)

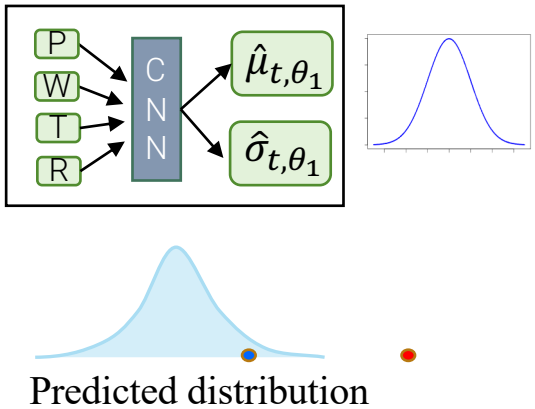
Step 2: Uncertainty informed (UI) anomaly score

Standard score



$$S_t^{(0)} = F(r_t; \mu^{(tr)}, \sigma^{(tr)})$$

UI score



$$\tilde{\mu}_t = \frac{1}{M} \sum_{m=1}^{M} \hat{s}_{t,\theta_m}$$

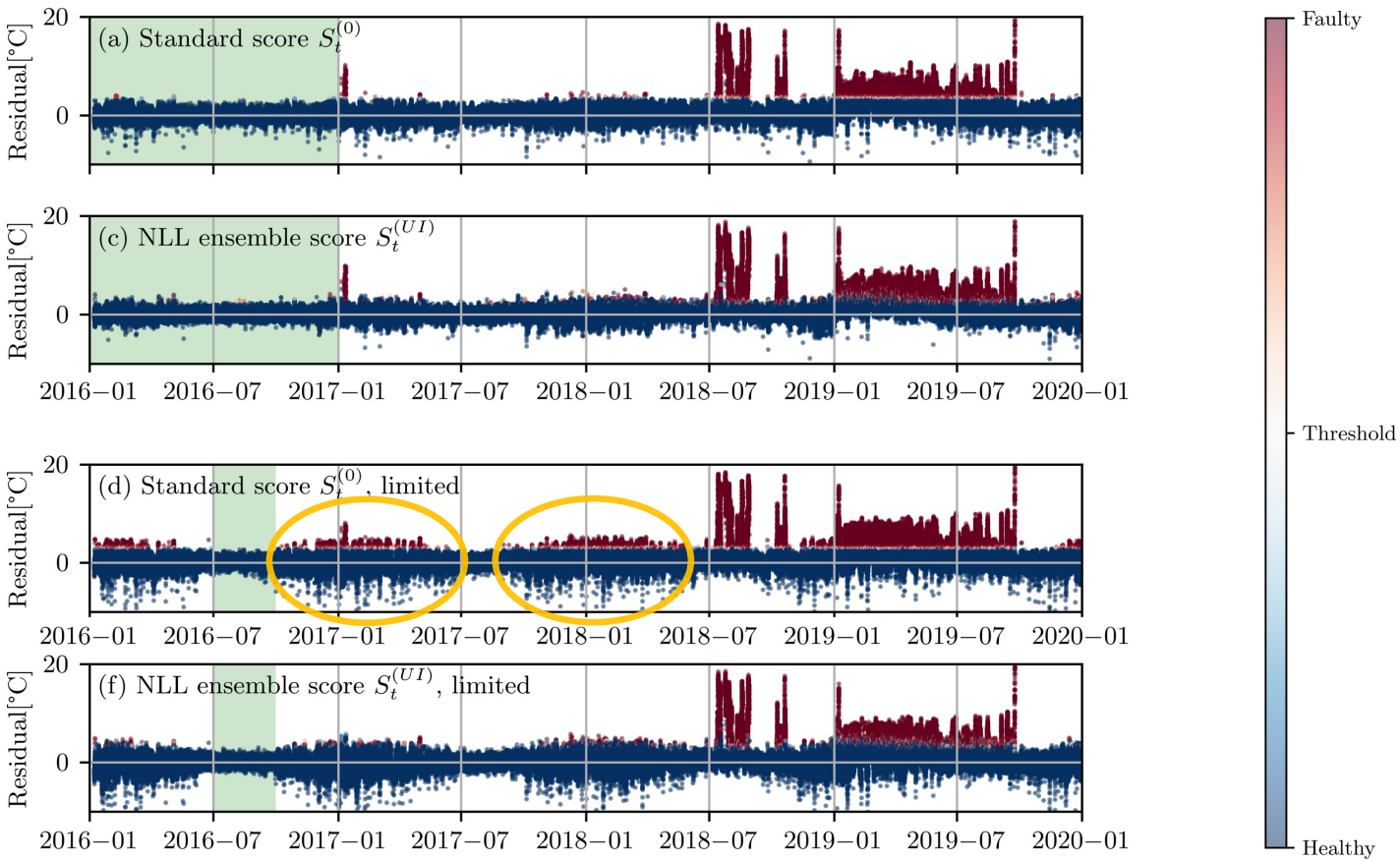
$$\tilde{\sigma}_t^2 = \frac{1}{M-1} \sum_{m=1}^{M} (\hat{s}_{t,\theta_m} - \tilde{\mu}_t)^2$$

$$S_t^{(UI)} = F(y_t; \tilde{\mu}_t, \tilde{\sigma}_t).$$

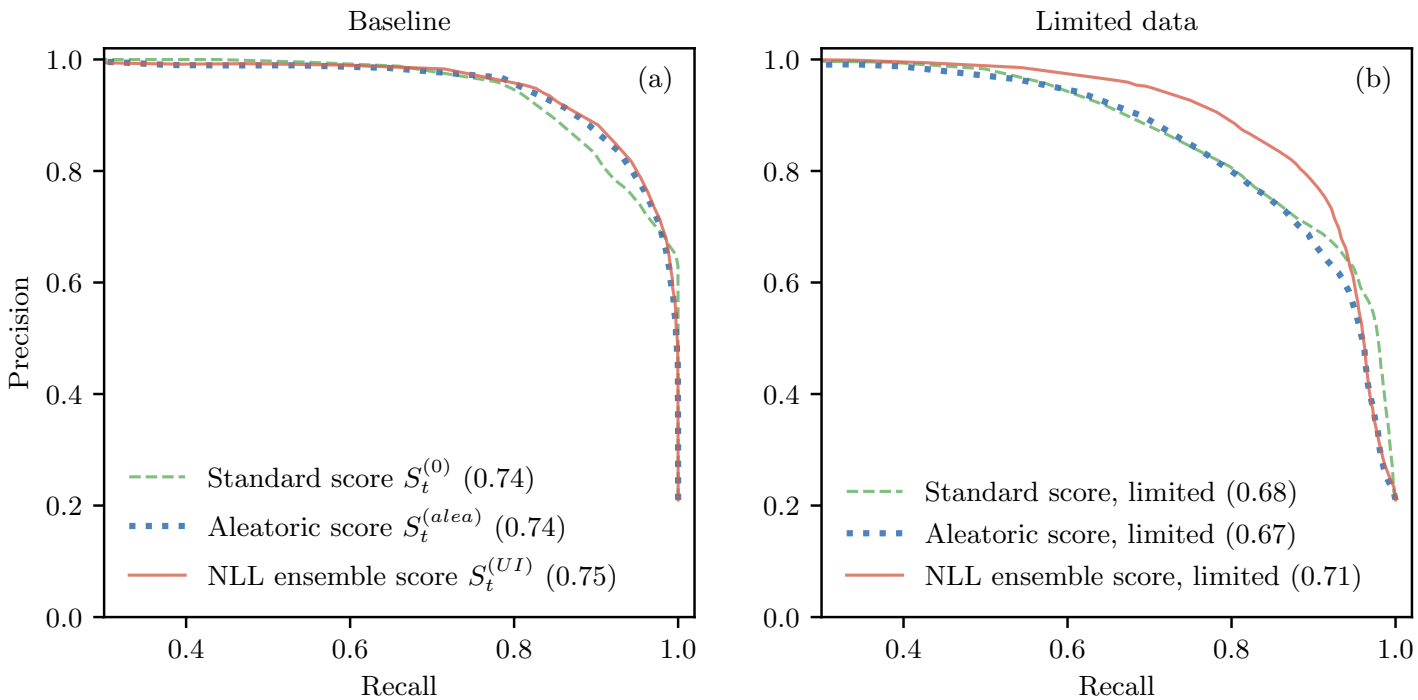
Anomaly detection evaluation

In distribution

Out of distribution



Anomaly detection: Precision-Recall

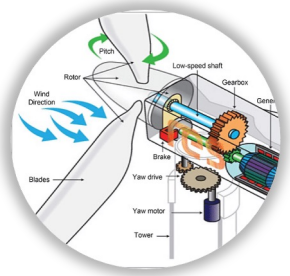


Zgraggen Jannik, et al. "Uncertainty informed anomaly scores with deep learning: robust fault detection with limited data" PHME Society, 2022.
BEST PAPER AWARD

Summary: generic concepts

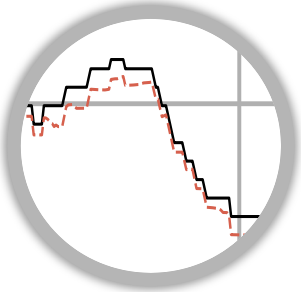
- Regression for anomaly detection **and localization**.
- Evaluate transfer learning using distribution shifts and reproducibility tests.
- Uncertainty informed anomaly scores.

AI in industry: collaboration of research and domain experts

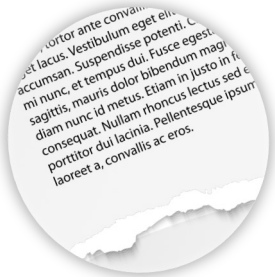


Ulmer, Markus, et al. "Early fault detection based on wind turbine scada data using convolutional neural networks." *5th European Conference of the Prognostics and Health Management Society,*. Vol. 5. No. 1. PHM Society, 2020.

Zgraggen, Jannik, et al. "Transfer Learning Approaches for Wind Turbine Fault Detection using Deep Learning". *6th European Conference of the Prognostics and Health Management Society,* Vol. 6. No. 1. PHM Society, 2021. ******Best Paper Award***



Zgraggen, Jannik, et al. "Uncertainty informed anomaly scores with deep learning: robust fault detection with limited data". *7th European Conference of the Prognostics and Health Management Society,* Vol. 7. No. 1. PHM Society, 2022. ******Best Paper Award***



Ulmer, Markus, et al. "Scaling up Deep Learning Based Predictive Maintenance for Commercial Machine Fleets: a Use Case", Proceedings of the SDS Conference 2022. ******Best Presentation Award**

Tutorial Notebooks

Contents

Contents