Fraunhofer IWES

Supported by:

Federal Ministry for Economic Affairs and Climate Action

on the basis of a decision by the German Bundestag

January 18, 2024 | EERA DeepWind 2024

Enhancing Wind Farm O&M with SCADA Data-Based Early Fault Detection

» Feasibility and Industry Expectations «

Timo Lichtenstein, Victor von Maltzahn

Industry Standpoint

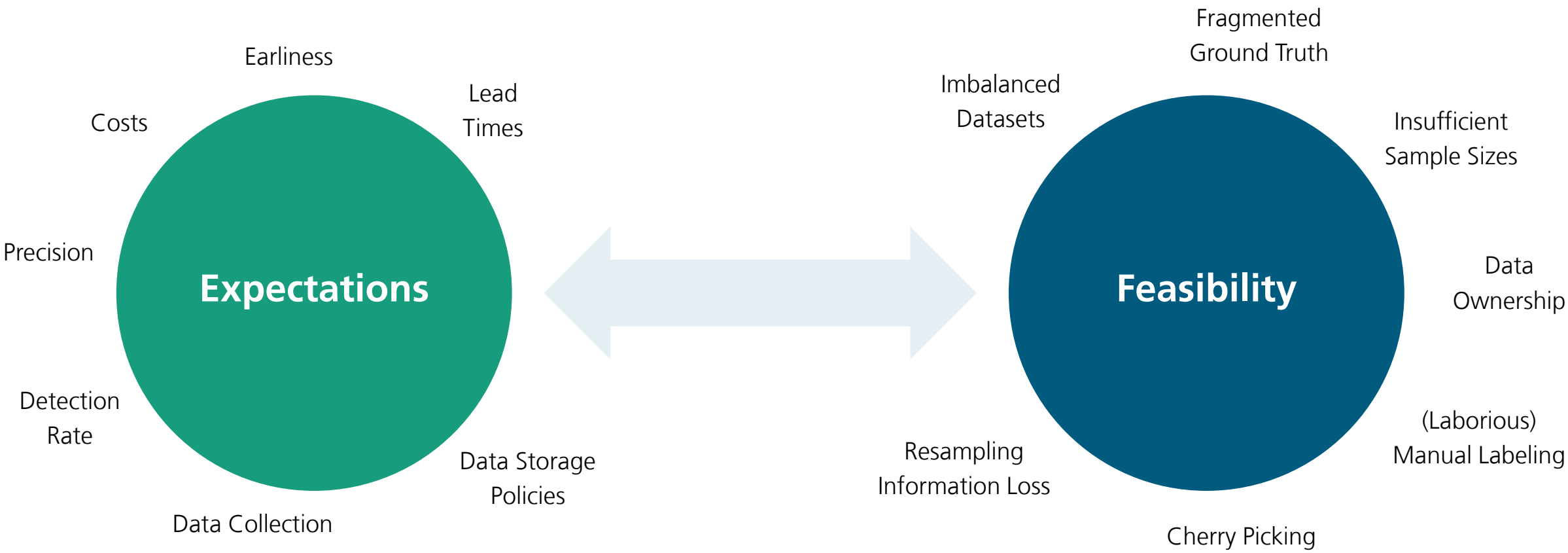
"The purpose of collecting reliability data for wind farms is not to identify the exact time or cause of a failure but rather to obtain information to allow operators and maintainers to intervene before failure or manage maintenance resources."

P. J. Tavner, *Offshore Wind Power – Reliability, availability and maintenance*, 2021

Fraunhofer IWES

Improving O&M With Early Fault Detection

What Is to Be Expected from Expectations?



Fraunhofer IWES

Understanding Industry Expectations

Method Development with Focus on Practical Applicability

Earliness

How early is a prediction needed?

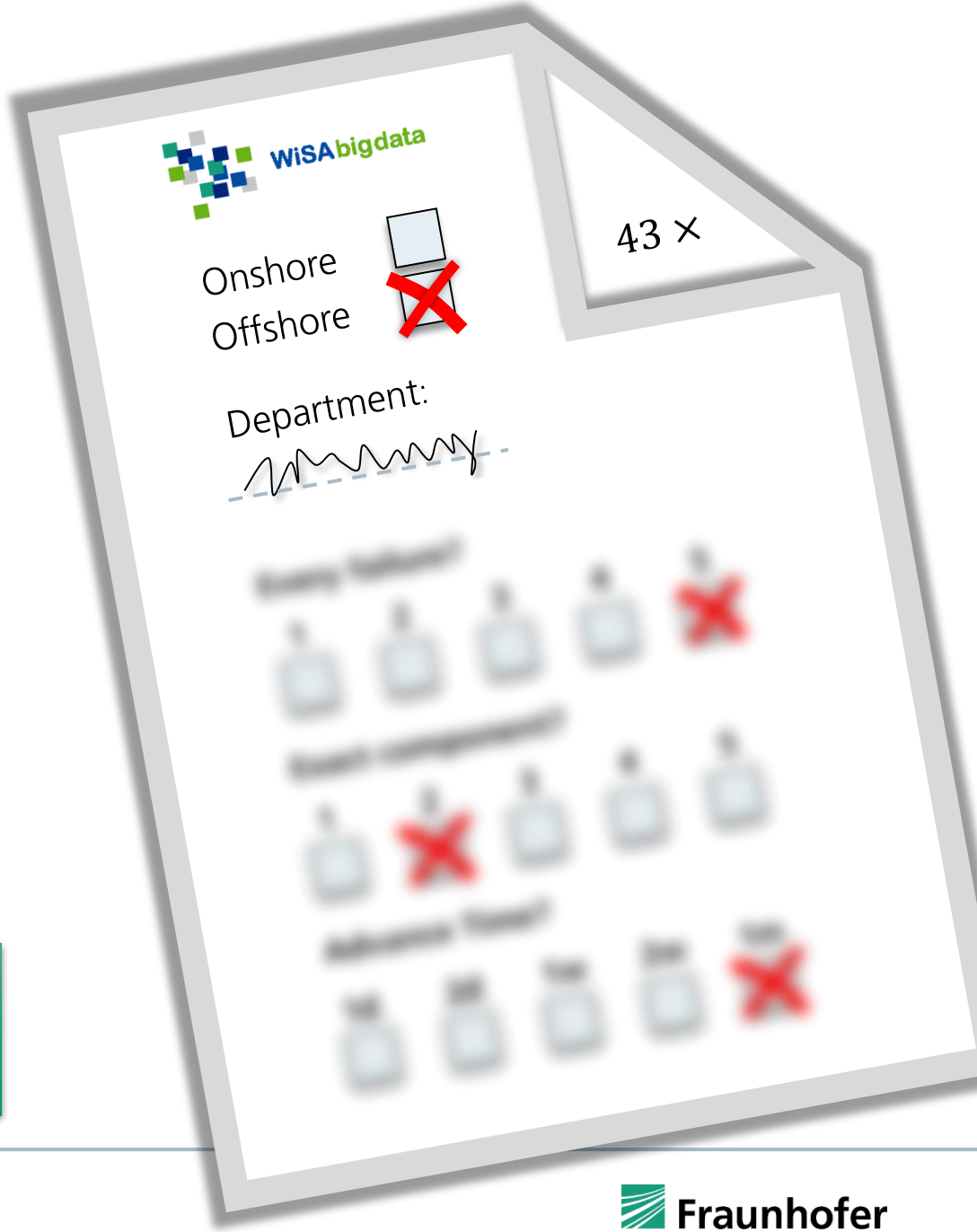
- Analysis of all lead times
- Desirable forecast times for process implementation

Performance

What is more important? Definite detection of all faults OR, in the case of a prediction, a certain fault detection?

- Analysis of severities and importance in processes
- Precision, Recall, F_β-Score

Ask the industry!



Fraunhofer IWES

Evaluation Metrics

Assessing Early Fault Detection Methods

Precision

Fraction of correctly detected faults among all predictions

$$P = \frac{TP}{TP + FP}$$

Recall

Fraction of correctly detected faults among all real faults

$$R = \frac{TP}{TP + FN}$$

F_β-Score

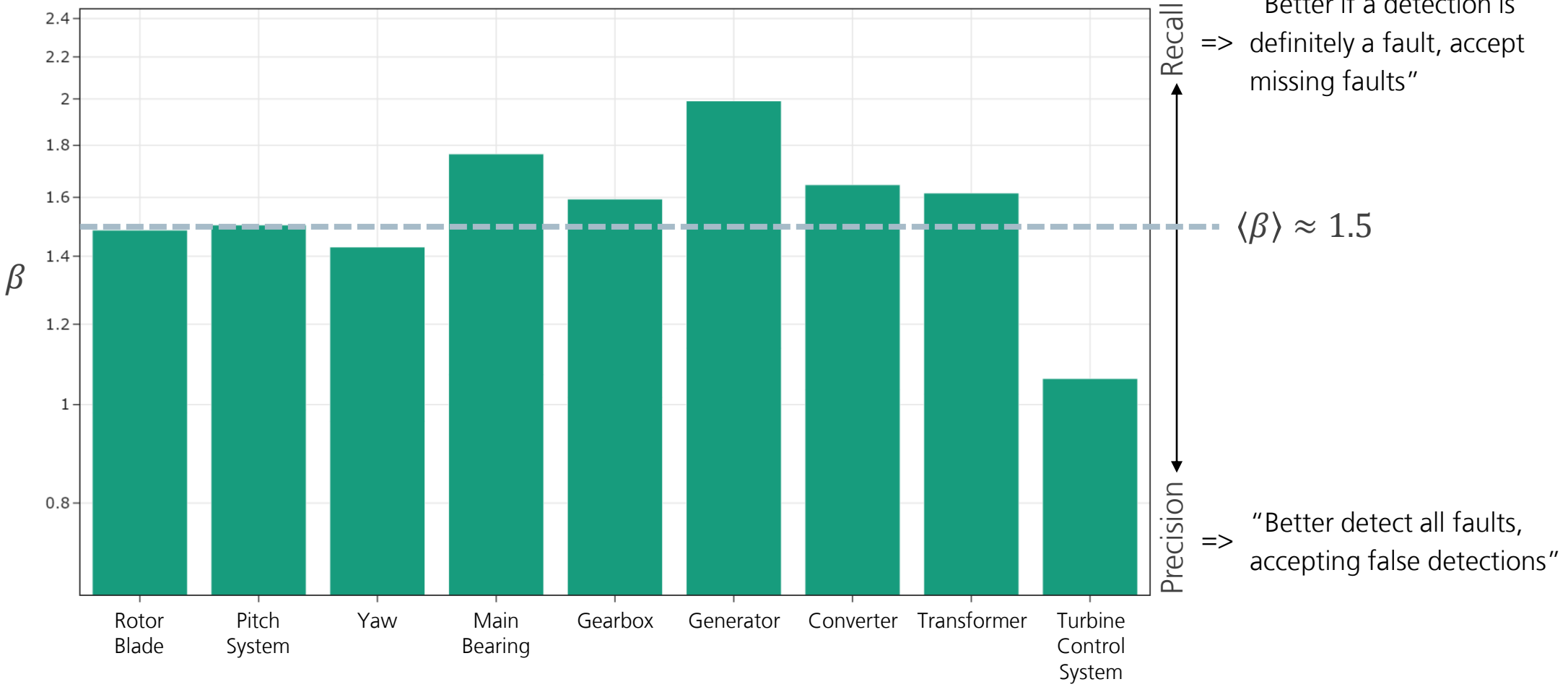
The weighted harmonic mean as a measure of performance

$$F_\beta = \frac{(1 + \beta^2)}{P^{-1} + \beta^2 \cdot R^{-1}}$$

"R is β times as important as P"

Expectations of the Industry – Performance Requirements

What is the Beta Weight for Each Component Category?



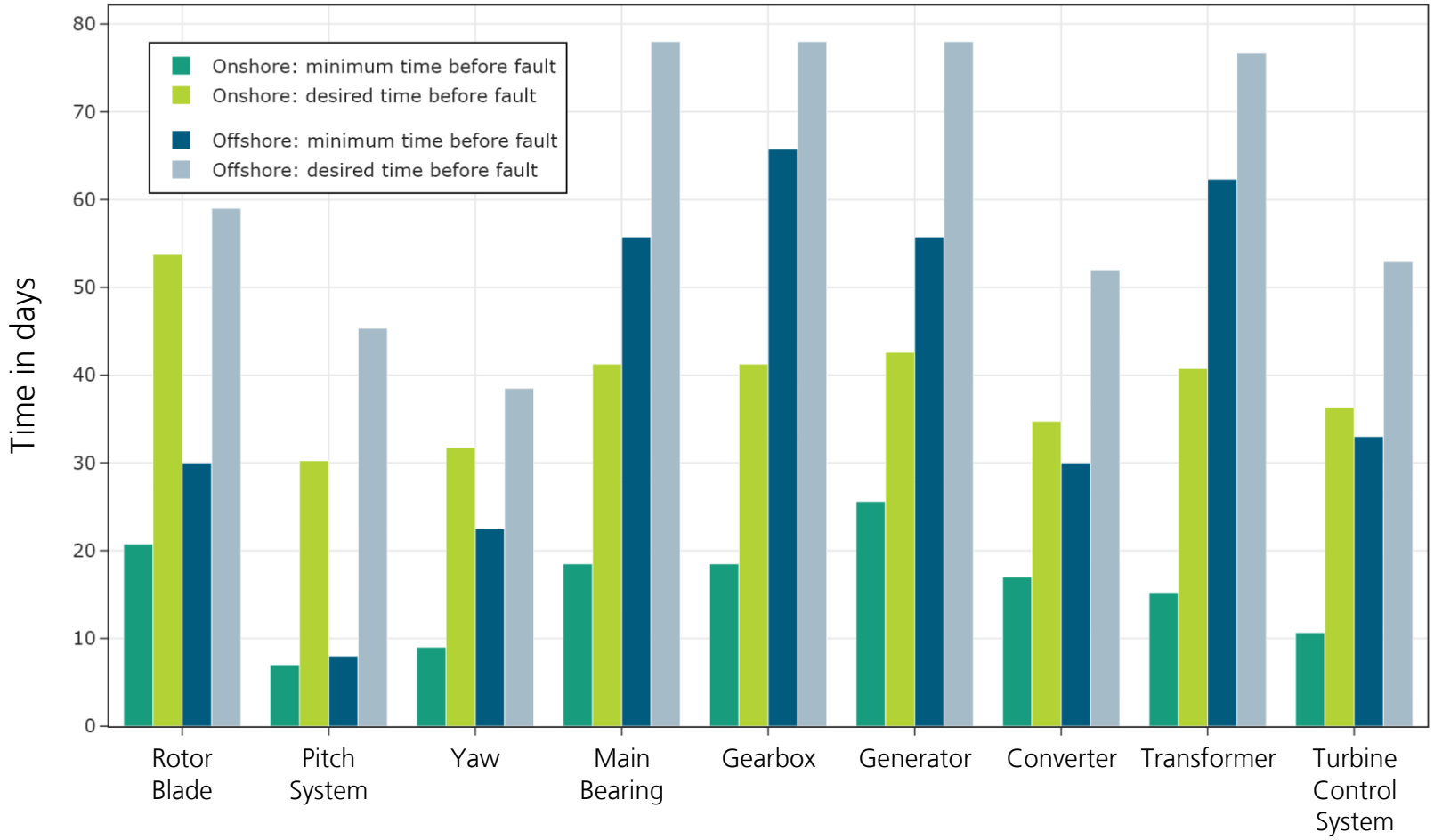
=> "Better if a detection is definitely a fault, accept missing faults"

=> "Better detect all faults, accepting false detections"

Fraunhofer IWES

Expectations of the Industry – Earliness Requirements

How Early Should an Early Fault Detection Method Give an Alarm?



Fraunhofer IWES

Challenges in SCADA Data EFD

Typical Obstacles in Research and Industry

Data Collection

- Resolution – data often only in 10 min resolution
- Ownership – access might be limited

Data Quality

- Varying data scope
  - SCADA operating data
  - SCADA alarm logs
  - Maintenance / service reports
- Company internal standards / naming schemes
- Data formats

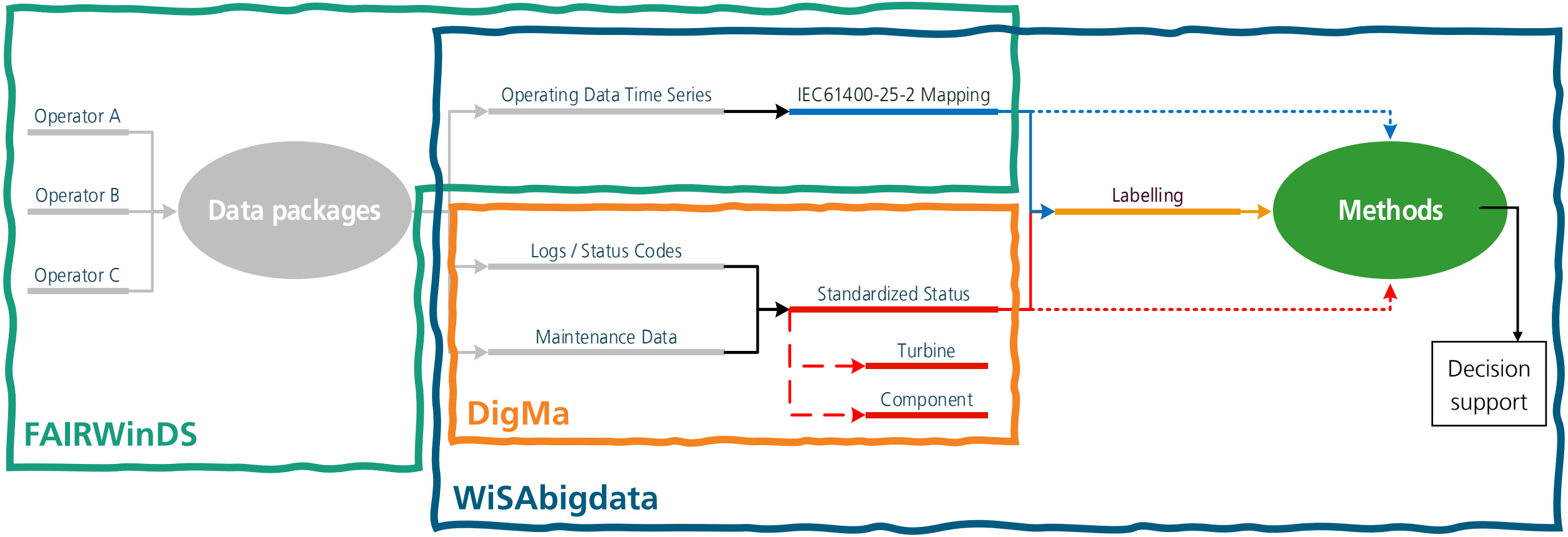
Data Handling

- Laborious preprocessing of maintenance data
- Definition of failures, faulty, and deviating states

Fraunhofer IWES

Simplified Data Value Creation Chain

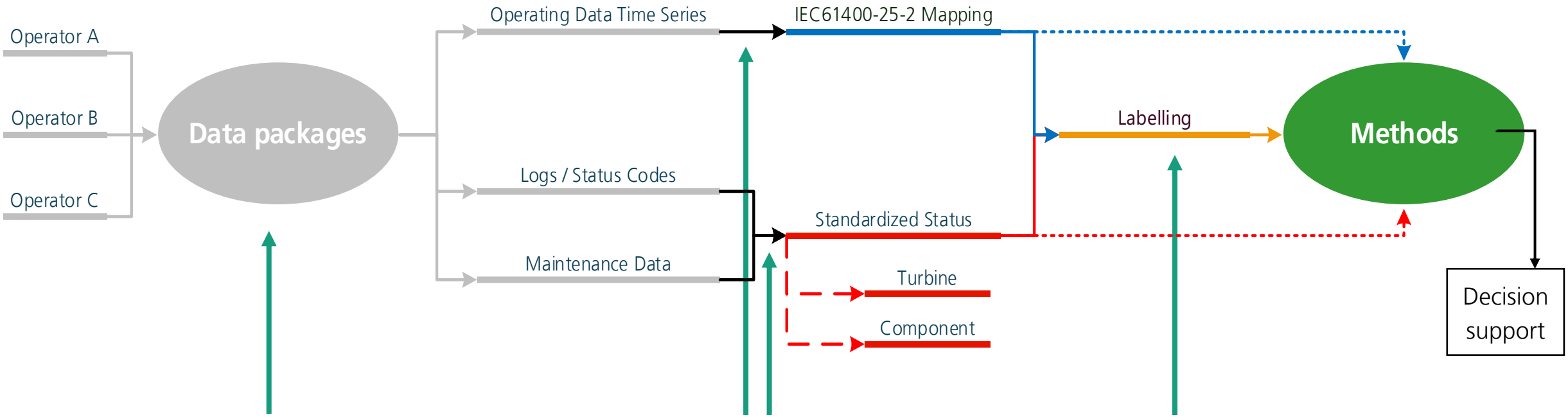
Example for Early Fault Detection with SCADA Data



Fraunhofer IWES

Simplified Data Value Creation Chain

Example for Early Fault Detection with SCADA Data



Data collection

- SCADA operating data from multiple turbines in varying resolutions
- SCADA alarm logs
- Free text service reports

Application of standards

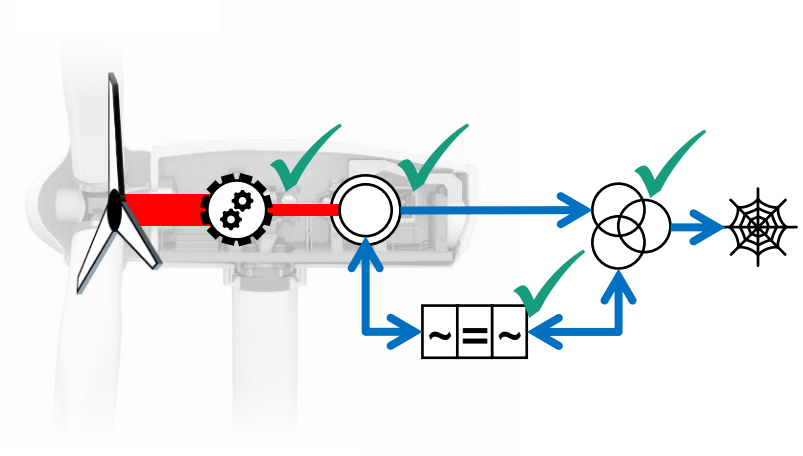
- Signal names: IEC61400-25-2
- Component names: RDS-PP
- States and events: ZEUS

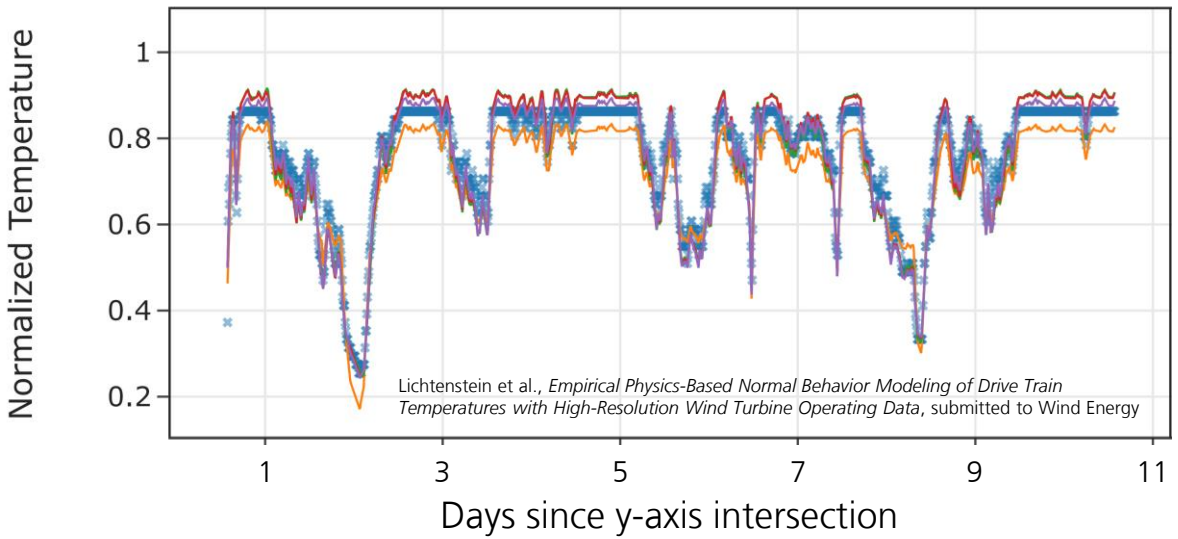
Deduction of ground truth

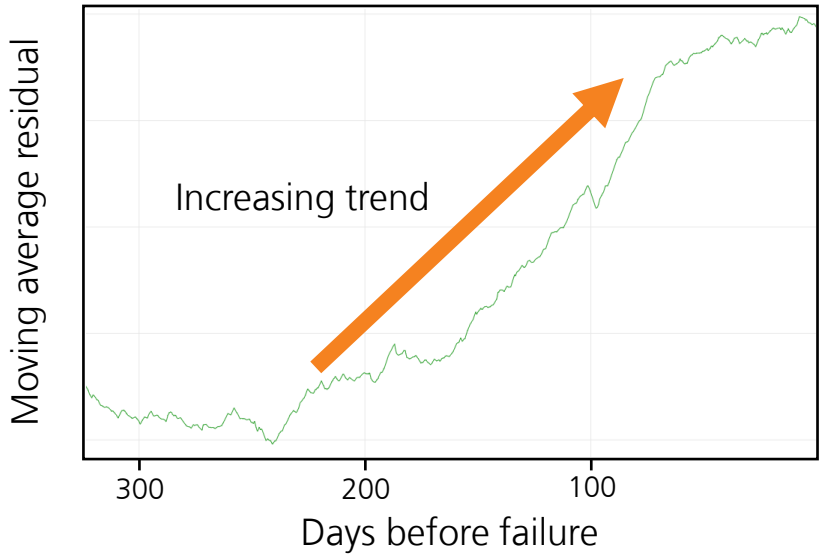
- Normal behavior and deviating operating conditions
- Failure := Component Replacements

Fault Detection Example

Using Normal Behavior Modeling







Example: Normal behavior models for main component temperatures

Multiple failures detectable, challenge to identify clear patterns

Current Challenges

For the Development of Early Fault Detection Using SCADA Data

Incomplete Datasets

- Partially incomplete data records, incorrect entries, incomprehensible comments, incorrect assignments
- Information hidden in non-disclosed datasets

Unclean Ground Truth Definitions

- Maintenance time is not necessarily failure time
  - Deriving the exact time of failure involves a great deal of effort
- Replacement not necessarily carried out due to impending failure
  - Classification problem in ground truth
- How to define a true detection and how early is a true detection a true detection?

Imbalanced Datasets

- Failures of certain components are very rare events
- All deviating states combined are a minority class

Improved digitalization in maintenance processes

Deal with missing data

Ground truth with uncertainty estimation

Methods independent of labels / ground truth

EU Data Act

(September 2025 at the latest)

Natural language processing algorithms

Summary and Recommendations for Action

For an Application of SCADA Data Early Fault Detection

Contact me:
Timo.Lichtenstein
@iwes.fraunhofer.de

Industry Expectations

- Recall is seen 1.5 times as important as precision in EFD
- Desired and indispensable prediction times for model development derived for every major component in on- and offshore maintenance

→ Industry-oriented assessment framework for method suitability

Feasibility

- Normal behavior modeling of turbine components
- Most model developments seem to be using “cherry picking”
- Current data quality and scope hinder classic pattern recognition and development of methods
- Approaches must stop relying on the improvement of the data basis

Improved digitalization in maintenance processes

Deal with missing data

→ Poster

Ground truth with uncertainty estimation

Methods independent of labels / ground truth

EU Data Act

(September 2025 at the latest)

Natural language processing algorithms

Fraunhofer IWES