Aalto-yliopisto
Perustieteiden
korkeakoulu

Mathematical Foundations of Binary Icing Classification of Wind Turbines

Markus Hakala

11.06.2026

Advisor: *Jonas Tölle*

Supervisor: *Jonas Tölle*

Aalto-yliopisto
Perustieteiden
korkeakoulu

Background

Why icing detection on wind turbines matters

1. **Energy loss**
Blade icing can cut annual energy production by over 20% in cold climates.

2. **Safety**
Ice shed from rotating blades endangers people and equipment nearby.

3. **Structural load**
Asymmetric ice causes vibration and accelerates fatigue of the structure.

4. **Maintenance & markets**
Early detection enables timely de-icing and better energy-market decisions.

Objectives

What the thesis aims to achieve

1. **Survey the methods**

Review peer-reviewed studies (mainly 2022–2025) for icing classification from SCADA data.

2. **Frame as binary classification**

Model as binary classification

3. **Mathematical Theory**

Present the mathematical foundations of the chosen classification method.

4. **Evaluate on real data**

Test how the method performs on wind-park SCADA data and report per-class metrics.

Scope

IN SCOPE

- SCADA-based classification of wind-turbine blade icing
- Binary classification: no icing / icing
- Methods from peer-reviewed literature, mainly 2022–2025
- Per-class evaluation (macro-F1, ROC-AUC)

OUT OF SCOPE

- Camera, ultrasonic or fibre-optic ice sensors
- Weather-forecast-only (NWP) models
- De-icing system design / control engineering

Sources and data

Literature base and the dataset used for testing

Literature

- 20 peer-reviewed studies (2018–2025, mainly 2022–2025)
- Databases: ScienceDirect, MDPI, Springer, Wiley, IOP

Test data

- SCADA data (Supervisory Control and Data Acquisition), from Databricks
- Features: power, wind speed, rotor RPM, acceleration, pitch, …
- Labelling: power deficit + ice-signal persistence rules

Methods and tools

Classical ML

Random Forest, XGBoost

Deep learning

CNN-Attention-GRU, MT-STAN, transformers

Semi-supervised

Tri-XGBoost, RFECV-TSVM, contrastive (UISSCL)

Tools

Python · scikit-learn · pandas · NumPy · Matplotlib

Schedule

1	Spring 2026	**Literature review**	Search, screen and synthesise peer-reviewed sources; finalise method landscape.
2	Summer 2026	**Method & theory**	Select the classification method and write its mathematical foundations.
3	Autumn 2026	**Writing**	Prepare SCADA data, train and evaluate the model, analyse per-class results.