POLITECNICO DI TORINO
Repository ISTITUZIONALE

Advanced Multivariate Anomaly Detection for Wind Turbines Using SCADA Data and Machine Learning

Original
Advanced Multivariate Anomaly Detection for Wind Turbines Using SCADA Data and Machine Learning / Daga, A.P., Fasana, A., Garibaldi, L., Gerbino, M., Viale, L.. - In: THE E-JOURNAL OF NONDESTRUCTIVE TESTING. - ISSN 1435-4934. - 2026-07:(2026). (14th European Conference on Non-Destructive Testing (ECNDT 2026) Verona (ITA) June 15-19, 2026) [10.58286/33271].

Availability:
This version is available at: 11583/3017171 since: 2026-10-28T18:34:31Z

Publisher:
NDT.net

Published
DOI:10.58286/33271

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

03 November 2026

e-Journal of Nondestructive Testing - ISSN 1435-4934 - www.ndt.net
14th European Conference on Non-Destructive Testing, Verona, Italy (ECNDT 2026)

Advanced Multivariate Anomaly Detection for Wind Turbines Using SCADA Data and Machine Learning

Alessandro Paolo DAGA [1], Alessandro FASANA [1], Luigi GARIBALDI [1], Marco GERBINO [1], and Luca VIALE [1,*]
[1] Department of Mechanical and Aerospace Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy
[*]*Corresponding author, E-Mail:* luca.viale@polito.it

More info about this article: https://www.ndt.net/?id=33271

1 Abstract

Ensuring the reliability of wind turbines is essential for reducing maintenance costs and improving energy availability. This study introduces a multivariate anomaly detection framework that combines physical modelling and machine learning to identify early signs of mechanical faults under variable operating conditions. The approach uses SCADA data and nacelle vibration measurements, integrating support vector regression. Unlike traditional residual-based methods, the proposed solution extends detection to a multivariate domain through Mahalanobis distance, improving robustness against confounding factors. Validation on an open-access offshore dataset demonstrates that the multivariate approach achieves high accuracy (95% on average) in detecting pitch-related failures and other anomalies, outperforming univariate residual analysis. These findings highlight the potential of diagnostic strategies for predictive maintenance in wind energy systems.

Keywords: wind turbines; SCADA data; anomaly detection; support vector regression; predictive maintenance; vibration monitoring.

2 Introduction

Wind turbines are complex systems operating under highly variable environmental conditions, which makes fault detection challenging [1,2]. Predictive maintenance strategies aim to anticipate failures before they lead to costly downtime [3,4]. SCADA systems provide continuous monitoring at low cost [5], offering large volumes of operational data, including vibration signals that are sensitive to mechanical degradation [6]. However, the high dimensionality and variability of these signals require advanced diagnostic methods [7]. This work proposes a hybrid approach that combines physical insights from aerodynamic modelling with machine learning techniques to improve anomaly detection performance.

The proposed approach employs Support Vector Regression (SVR) trained on healthy operational data to model the normal relationship between turbine conditions and vibration levels. This predictive model enables the estimation of expected vibration under given operating states, so that deviations from the prediction indicate potential faults rather than normal variations due to wind or load. The framework extends this principle to a multivariate setting by combining multiple indicators and statistical techniques, improving robustness and supporting early detection of mechanical anomalies.

3 Methodology

The proposed framework focuses on multivariate anomaly detection and fault classification using SCADA data and vibration measurements. The process includes:

- *Data Acquisition and Preprocessing:* SCADA data are cleaned and filtered to remove outliers. Relevant features include wind speed, pitch angle, rotor speed, gearbox oil temperature, and nacelle vibration signals. Dimensionality reduction is performed using PCA to capture the most informative components.
- *Physical Modelling:* Aerodynamic power estimation based on Betz theory is

© 2026 The Authors. Published by NDT.net under License CC-BY-4.0 https://creativecommons.org/licenses/by/4.0/
https://doi.org/10.58286/33271

incorporated to improve interpretability and account for operational variability.

- *Predictive Modelling:* Multiple SVR models are trained on healthy operational data to predict vibration indicators under normal conditions. This step establishes a reference model that accounts for the influence of operating conditions [8,9].
- *Multivariate Anomaly Detection:* Deviations between predicted and measured values (residuals) are used to compute an anomaly index as a Mahalanobis distance.
- *Fault Classification:* To extend beyond novelty detection, classification is applied to recognize fault types and enable multi-damage identification.

4 Results

The multivariate approach based on Mahalanobis distance achieved high detection performance, with ROC-AUC and accuracy values exceeding 0.93 and 94%, respectively. This confirms the robustness of the method under varying operational conditions. Classification experiments demonstrated that pitch-related failures and other mechanical anomalies can be distinguished with accuracy above 95%, supporting the feasibility of multi-class fault identification. These results indicate that the proposed framework not only detects anomalies but also provides a basis for fault categorization, enhancing predictive maintenance strategies.

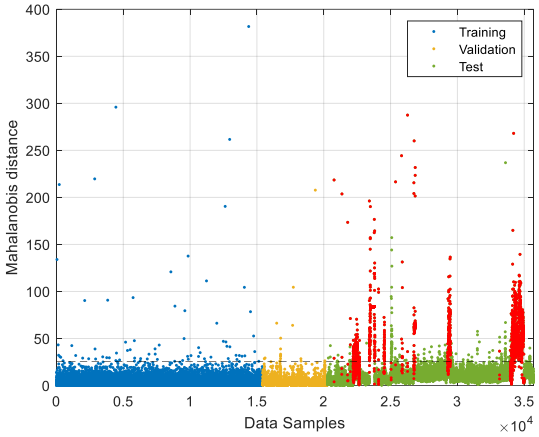


Fig. 1. Mahalanobis distance. Red markers denote the labelled fault data of WT35

5 Conclusions

The integration of physical modelling and machine learning offers a powerful tool for anomaly detection in wind turbines. While residual-based methods provide a baseline solution, multivariate strategies deliver superior accuracy and robustness, enabling early fault identification without interrupting turbine operation. Future work may focus on improving accuracy, enhancing model generalization across different turbine types, and exploring alternative learning architectures to support scalability and real-time deployment in large wind farms.

References

1. R.S. Amano, Review of Wind Turbine Research in 21 Cent., J. En. Res. Tech. (2017).
2. R. Kumar, et al., Damage detection of wind turbine system based on signal processing approach: a critical review, Clean Technol. Environ. Policy 23 (2021) 561–580.
3. L. Viale, et al., Condition monitoring of an industrial book-cutting machine using a novel mixture of vibration and physics-based features, Int. J. Comput. Integr. Manuf. (2025).
4. L. Viale, et al., Data-Driven Fault Diagnosis for Rotating Industrial Paper-Cutting Machinery, Machines 13 (2025) 688.
5. J. Tautz-Weinert, et al., Using SCADA data for wind turbine condition monitoring – a review, IET Renew. Power Gener. 11 (2017) 382–394.
6. F. Castellani, et al., Analyzing wind turbine directional behavior: SCADA data mining techniques for efficiency and power assessment, Appl. Energy 185 (2017) 1076–1086.
7. L. Viale, et al., Dimensionality Reduction Methods of a Clustered Dataset for the Diagnosis of a SCADA-Equipped Complex Machine, Machines 11 (2023) 36.
8. M.A. Pimentel, et al., A review of novelty detection, Signal Proc. 99 (2014) 215–249.
9. K. Worden, Structural Fault Detection Using A Novelty Measure, J. S. Vib. 201 (1997).