REPUBLIC OF TURKEY
ISTANBUL GELISIM UNIVERSITY
INSTITUTE OF GRADUATE STUDIES

Department of Electrical and Electronics Engineering

ARTIFICIAL INTELLIGENCE APPLICATIONS IN MONITORING AND MAINTAINING WIND TURBINES

Master Thesis

Zeyad Tareq Ahmed AHMED

Supervisor

Asst. Prof. Dr. Ercan AYKUT

Istanbul – 2026

THESIS INTRODUCTION FORM

Name and Surname : Zeyad Tareq Ahmed AHMED

Language of the Thesis : English

Name of the Thesis : Artificial Intelligence Applications in Monitoring and Maintaining Wind Turbines

Institute : Istanbul Gelisim University Institute of Graduate Studies

Department : Electrical and Electronics Engineering

Thesis Type : Master

Date of the Thesis : 18 / 06 /2026

Page Number : 98

Thesis Supervisor : Asst. Prof. Dr. Ercan AYKUT

Index Terms : wind turbines; fault detection; deep learning

Turkish Abstract : Bu araştırma, sistem güvenilirliğini artırmak ve bakım maliyetlerini azaltmak amacıyla fiziksel modellemeyi yapay zeka teknikleriyle birleştirerek rüzgar türbini arızalarının erken teşhisi için entegre bir akıllı çerçeve sunmaktadır. Bu çalışmanın önemi, gerçek dünya verilerinin eksikliği ve geleneksel modellerin karmaşık çalışma ortamlarında genelleme yapma konusundaki sınırlı yetenekleriyle ilgili zorlukları ele almasında yatmaktadır. Bunu başarmak için, çeşitli çalışma koşulları altında termal ve mekanik davranışı temsil eden sinyaller üretmek üzere MATLAB/Simulink kullanılarak dinamik bir model geliştirilmiştir. Ardından SVM, ANN, Otomatik Kodlayıcı ve hibrit CNN–LSTM modeli dahil olmak üzere çeşitli yapay zeka modelleri uygulanmıştır. Sonuçlar,

CNN–LSTM hibrit modelinin doğruluk açısından üstünlüğünü ortaya koymuş; model, tüm arıza durumlarını tespit etme konusundaki yüksek yeteneğini yansıtan %99,84'lük bir doğruluk oranı ve 1,0000'lık bir geri çağırma değeri elde etmiştir. Bununla birlikte, sonuçlar hassasiyette nispi bir düşüş göstermiş ve bu da bazı yanlış alarmların varlığına işaret etmiştir. Bu araştırma, öngörücü bakım çalışmalarını destekleyebilecek simülasyon temelli bir çerçeve sunmaktadır. Bununla birlikte, pratik rüzgâr türbini sistemlerinde uygulanabilirliğini doğrulamak için gerçek dünya verileriyle daha kapsamlı bir doğrulama çalışması gereklidir.

Distribution List : 1. To the Institute of Graduate Studies of Istanbul Gelisim University
2. To the National Thesis Center of YÖK (Higher Education Council)

Signature

Zeyad Tareq Ahmed AHMED

REPUBLIC OF TURKEY
ISTANBUL GELISIM UNIVERSITY
INSTITUTE OF GRADUATE STUDIES

Department of Electrical and Electronics Engineering

ARTIFICIAL INTELLIGENCE APPLICATIONS IN MONITORING AND MAINTAINING WIND TURBINES

Master Thesis

Zeyad Tareq Ahmed AHMED

Supervisor

Asst. Prof. Dr. Ercan AYKUT

Istanbul – 2026

DECLARATION

I hereby declare that in the preparation of this thesis, scientific ethical rules were followed, the works of other people were referenced in accordance with scientific norms, if used, there was no falsification in the used data, and no part of the thesis was submitted to this university or any other university as another thesis.

Zeyad Tareq Ahmed AHMED

.... /.... /2026

TO ISTANBUL GELISIM UNIVERSITY

THE DIRECTORATE OF INSTITUTE OF GRADUATE STUDIES

The thesis study of Zeyad Tareq Ahmed Ahmed titled as Artificial Intelligence Applications in Monitoring and Maintaining Wind Turbines has been accepted as MASTER THESIS in the department of Electrical and Electronics Engineering by out jury.

Signature

Director — *Asst. Prof. Dr. Ercan AYKUT*
(Supervisor)

Signature

Member — *Asst. Prof. Dr. Halit YAHYA*

Signature

Member — *Asst. Prof. Dr. Ulvi BAŞPINAR*

APPROVAL

I approve that the signatures above signatures belong to the aforementioned faculty members.

... / ... / 2026

Signature

Prof. Dr. İzzet GÜMÜŞ

Director of the Institute

SUMMARY

This research presents an integrated intelligent framework for the early detection of wind turbine failures by combining physical modeling with artificial intelligence techniques, with the aim of improving system reliability and reducing maintenance costs. The significance of this work lies in addressing the challenges associated with a lack of real world data and the limited ability of traditional models to generalize in complex operating environments. To achieve this, a dynamic model was developed using MATLAB/Simulink to generate signals representing thermal and mechanical behavior under various operating conditions. Several AI models were then applied, including SVM, ANN, Autoencoder, and the hybrid CNN-LSTM model. The results demonstrated the superiority of the CNN-LSTM hybrid model in terms of accuracy, achieving an ac-curacy of 99.84% with a recall value of 1.0000, reflecting its high ability to detect all failure cases. However, the results showed a relative decrease in precision, indicating the presence of some false alarms. This research provides a simulation based framework that can support predictive maintenance research. However, further validation using real world data is required to confirm its applicability in practical wind turbine systems.

Key Words: wind turbines; fault detection; deep learning

ÖZET

Bu araştırma, sistem güvenilirliğini artırmak ve bakım maliyetlerini azaltmak amacıyla fiziksel modellemeyi yapay zeka teknikleriyle birleştirerek rüzgar türbini arızalarının erken teşhisi için entegre bir akıllı çerçeve sunmaktadır. Bu çalışmanın önemi, gerçek dünya verilerinin eksikliği ve geleneksel modellerin karmaşık çalışma ortamlarında genelleme yapma konusundaki sınırlı yetenekleriyle ilgili zorlukları ele almasında yatmaktadır. Bunu başarmak için, çeşitli çalışma koşulları altında termal ve mekanik davranışı temsil eden sinyaller üretmek üzere MATLAB/Simulink kullanılarak dinamik bir model geliştirilmiştir. Ardından SVM, ANN, Otomatik Kodlayıcı ve hibrit CNN–LSTM modeli dahil olmak üzere çeşitli yapay zeka modelleri uygulanmıştır. Sonuçlar, CNN–LSTM hibrit modelinin doğruluk açısından üstünlüğünü ortaya koymuş; model, tüm arıza durumlarını tespit etme konusundaki yüksek yeteneğini yansıtan %99,84'lük bir doğruluk oranı ve 1,0000'lık bir geri çağırma değeri elde etmiştir. Bununla birlikte, sonuçlar hassasiyette nispi bir düşüş göstermiş ve bu da bazı yanlış alarmların varlığına işaret etmiştir. Bu araştırma, öngörücü bakım çalışmalarını destekleyebilecek simülasyon temelli bir çerçeve sunmaktadır. Bununla birlikte, pratik rüzgâr türbini sistemlerinde uygulanabilirliğini doğrulamak için gerçek dünya verileriyle daha kapsamlı bir doğrulama çalışması gereklidir.

Anahtar kelimeler: rüzgar türbinleri; arıza tespiti; derin öğrenme

TABLE OF CONTENTS

SUMMARY ..... **i**
ÖZET ..... **ii**
TABLE OF CONTENTS ..... **iii**
ABBREDIVATIONS ..... **v**
LIST OF TABLES ..... **vii**
LIST OF FIGURES ..... **viii**
PREFACE ..... **x**

CHAPTER ONE
GENERAL INTRODUCTION

1.1 Introduction 1
1.2 Motivation 4
1.3 Problem Statement 4
1.4 Aims 6
1.5 Thesis Contributions 6
1.6 Thesis Outlines 7

CHAPTER TWO
LİTERATURE REVIEW

2.1 Introduction 9
2.2 Overview of the Wind Turbine System 9
2.2.1 Traditional Monitoring Methods 10
2.2.2 SCADA-based monitoring 11
2.2.3 Data-Driven Monitoring 12
2.2.4 Conclusion 12
2.3 SCADA-Based Fault Detection 12
2.3.1 Introduction to SCADA in Wind Turbines 12
2.3.2 Fault detection mechanism using SCADA data 14
2.3.3 Challenges and limitations in SCADA systems 16
2.3.4 Recent developments in integrating SCADA systems with artificial intelligence 17
2.3.5 Conclusion 19

CHAPTER THREE
METHODOLOGY AND SYSTEM DESIGN

3.1 Introduction 20
3.2 Overall Research Methodology 20
3.3 Wind Turbine System Modeling in Simulink 22
3.3.1 Wind Turbine System Modeling in Simulink 22

3.3.2 Generator and Electrical Subsystem 24
3.3.3 Control Strategy 24
3.3.4 Integration with Thermal and Vibration Systems 24
3.4 Sensing and Monitoring Subsystem 25
3.5 Proposed AI Model 28
3.6 Implementation Process 34
3.7 Performance Evaluation 37
3.8 Conclusion 38

CHAPTER FOUR
SIMULATION AND SYSTEM IMPLEMENTATION

4.1 Introduction 39
4.2 Simulation Environment & Configuration 39
4.3 Wind Turbine Model in Simulink 40
4.4 Fault Injection 41
4.5 Data Acquisition 43
4.6 Data Preparation for Artificial Intelligence 43
4.7 Summary 44

CHAPTER FIVE
RESULTS AND AI-BASED FAULT DIAGNOSIS

5.1 Introduction 46
5.2 Simulation Results 47
5.2.1 Normal Operation 49
5.2.2 Fault Operation 49
5.3 Feature Analysis 50
5.3.1 Introduction to Extracted Features 50
5.3.2 Analysis of characteristics on day 10 55
5.3.3 Analysis of characteristics on day 56 55
5.3.4 Analysis of characteristics on day 208 56
5.3.5 General discussion and engineering implications 56
5.4 Results of artificial intelligence models for fault detection 57
5.4.1 Autoencoder-Based Anomaly Detection 58
5.4.2 SVM-Based Anomaly Classification 61
5.4.3 Artificial Neural Network (ANN) Results 63
5.4.4 Hybrid CNN–LSTM Anomaly Detection Model 66
5.5 Overall Comparative Discussion and Final Remarks 69

CONCLUSIONS AND RECOMMENDATIONS 73
REFERENCES 75
RESUME 80

ABBREDIVATIONS

AI	:	Artificial Intelligence
ANN	:	Artificial Neural Networks
AUC	:	Area Under the Curve
AutoML	:	Automated Machine Learning
Bi-LSTM	:	Bidirectional Long Short-Term Memory
CMS	:	Condition Monitoring System
CNN	:	Convolutional Neural Network
DC	:	Direct Current
DFIG	:	Doubly Fed Induction Generator
GAN	:	Generative Adversarial Network
ICA	:	Independent Component Analysis
KNN	:	K-Nearest Neighbor
LSTM	:	Long Short-Term Memory
MRNet	:	Manitoba's Regional Advanced Research and Education Network
MAT	:	Matlab Workspace Files
MLP	:	Multi-Layer Perceptron
MPPT	:	Maximum Power Point Tracking
MSE	:	Mean Squared Error
P2P	:	Peak-To-Peak
PCA	:	Principal Component Analysis
PDE	:	Partial Differential Equation
PI	:	Proportionally-Integral
PLC	:	Programmable Logic Controller
PMSG	:	Permanent Magnet Synchronous Generator
PMSM	:	Permanent Magnet Synchronous Motor
PS-Simulink	:	Converts A Physical Signal into a Simulink® Output Signal
RBF	:	Radial Basis Function
ReLU	:	Rectified Linear Unit
RF	:	Random Forest
RMS	:	Root Mean Square

RMSE	:	Root Mean Square Error
RNN	:	Recurrent Neural Networks
ROC	:	Receiver-Operating Characteristic Curve
SCADA	:	Supervisory Control and Data Acquisition
SMC	:	System Management Controller
STD	:	Standard Deviation
SVM	:	Support Vector Machine
XGBoost	:	Extreme Gradient Boosting

LIST OF TABLES

Table 1 Wind Turbine Simulation and Training Parameters..... 23
Table 2 Hyperparameters of the Proposed AI Models. 52
Table 3 Statistical characteristics of the (Heat Signal)..... 53
Table 4 Statistical Properties of (Vibration Signal) 54
Table 5 Quantitative analyses of the reconstruction error sensitivity of the CNN–LSTM model. 58
Table 6 Performance metrics of Autoencoder anomaly detection model 60
Table 7 Performance metrics of SVM-based anomaly detection model..... 63
Table 8 Performance metrics of ANN anomaly detection model 65
Table 9 Performance metrics of Hybrid CNN–LSTM anomaly detection model 68
Table 10 Comparison of selected studies on wind turbine fault diagnosis. 70

LIST OF FIGURES

Figure 1 Flow chart of the proposed method for wind-turbine fault detection (adapted from (J. Liu et al., 2023)). 5
Figure 2 Temperature-based condition monitoring framework for wind turbines ((Al Lahham et al., 2025) 11
Figure 3 SCADA monitoring techniques in wind turbines ((Al Lahham et al., 2025) 13
Figure 4 Fault detection mechanism using SCADA data (regression decision logic) ((Sheng et al., 2025), Neurocomputing, 633, 129830, Fig. 10) 15
Figure 5 Two-layer framework for integrating SCADA data with artificial intelligence in wind turbine systems (Behara & Saha, 2025a) 18
Figure 6 Overall methodology steps 21
Figure 7 Complete physical model of a wind turbine in the Simulink environment, showing the turbine, generator, power converter, and control and measurement system units. 25
Figure 8 Data acquisition and processing workflow for temperature and vibration signals in the wind. 27
Figure 9 Internal structure of the LSTM algorithm used in analyzing heat and vibration signals to predict abnormal conditions in wind turbines (Nabi et al., 2021). 30
Figure 10 The general structure of the Convolutional Neural Network (CNN) used in classifying the operating conditions of wind turbines (Zare & Ayati, 2021). 31
Figure 11 Simplified diagram of an ANN network for classifying the operating status of a wind turbine 32
Figure 12 The conceptual structure of the SVM algorithm in classifying the operating conditions of wind turbines 33
Figure 13 The proposed structure of the Autoencoder model used to detect anomalies in the temperature and vibration data of wind turbine 34
Figure 14 Diagram of the steps for implementing the proposed diagnostic system in the MATLAB/Simulink environment 36
Figure 15 Response of thermal and vibration signals during fault injection in a physical model of a wind turbine. 42
Figure 16 Raw wind signal before clipping (Istanbul – 2024) 47
Figure 17 Wind signal after clipping to the operating range (Safe Operating Zone: 3–25 m/s) 47
Figure 18 Temperature signal during normal operation (Normal Temperature) 48
Figure 19 Temperature signal during a fault condition (Fault Temperature – Drift/Instability) 48
Figure 20 Normal Vibration Signal 49
Figure 21 Wind signal during fault (Fault @ Day 56) 50
Figure 22 Comparison of heat and vibration characteristics for days (10,56,208) ... 53

Figure 23 The convergence behavior of the CNN–LSTM model is illustrated by the decrease in RMSE and loss values during training. 57
Figure 24 shows the confusion matrix results for the Autoencoder model for days (10, 56, 208) 59
Figure 25 shows the ROC curves for the Autoencoder model for the three days 60
Figure 26 shows the confusion matrices for the SVM model for days (10, 56, 208) 61
Figure 27 shows the ROC curves for the SVM model for the three days 62
Figure 28 shows the confusion matrices for the ANN model for days (10, 56, 208) 64
Figure 29 shows the ROC curves for the ANN model for the three days. 65
Figure 30 shows the confusion matrices for the Hybrid CNN–LSTM model for days (10, 56, 208) 66
Figure 31 shows the ROC curves for the Hybrid CNN–LSTM model for the three days is same 67
Figure 32 Comparison of fault detection algorithms on the test dataset 69

PREFACE

I praise Allah, the Almighty, in the first place, with praise befitting His majesty, for His guidance, patience, and blessings upon me, and for the strength that enabled me to complete this thesis.

I would also like to express my sincere thanks and gratitude to my family and supporters for their continuous support, constant encouragement, and patience throughout my academic career, as they had the greatest impact on my perseverance and completion of this work. I would also like to express my sincere thanks and appreciation to my thesis supervisor Asst. Prof. Dr. Ercan Aykut for his scientific guidance, careful follow-up, and valuable comments, which contributed to the development of this thesis and the improvement of its scientific level. I would like to express my gratitude to the members of the discussion committee for their time and effort, and for their constructive scientific comments and suggestions, which contributed to improving the quality of this research work. I would also like to thank the faculty at the university and the Institute of Science for providing a suitable scientific environment and resources that helped me complete this study.

CHAPTER ONE

GENERAL INTRODUCTION

1.1 Introduction

Wind energy systems offer high performance among renewable energies and are one of the fastest growing types of renewable energy in the world, providing a solid foundation for long-term commitment to clean energy initiatives and carbon emission reduction (Aslam et al., 2021; Qin et al., 2023a; Wadawa & Effa, 2025). With the expansion of wind farms and the increase in turbine capacity in challenging onshore and offshore conditions, operational reliability and continuity are important factors in ensuring economic consequences (Xiang et al., 2021; Ohunakin et al., 2024). Today's turbines are highly complex machines containing numerous mechanical and electrical systems that suffer from cyclic stresses and thermal changes that cause slow failures that are not easily detectable using traditional methods (Bouthaina Abichou et al., 2014; Ghanbari et al., 2023; Mulongo et al., 2020). The pitch system is one of the most sensitive systems (the system responsible for managing blade angles), and unfavorable positioning can lead to significant losses in energy efficiency and increased maintenance costs (Binti Shahrulhisham et al., 2022) . With the advancement of technology, diagnostic systems such as SCADA have become indispensable for collecting and analyzing data from multiple points on the turbine (Sheng et al., 2025). These units capture tens of thousands of measurements in real time, enabling the application of big data to analyze health status and identify deviations in turbine performance (Sheng et al., 2025).

However, SCADA data is very complex as it is high-dimensional, unbalanced (the numbers corresponding to failure cases occupy only a very small proportion of the whole), and often noisy or incomplete (Pereira & Silveira, 2018; Sheng et al., 2025). This complexity also leads to limitations in these analyses using statistical and traditional models(Alagha et al., 2025a; Rossi et al., 2016). Accordingly, the current trend is to abandon physical models and focus on machine learning and deep learning algorithms to take into account the complexity of the data structure and improve diagnostic accuracy (Alagha et al., 2025a). Support vector machine (SVM), random forest (RF), and k-nearest neighbor (KNN) algorithms are commonly used methods

for analyzing SCADA data to detect common faults in the industry. Despite their simplicity and success in some cases, they require balanced and accurately labeled data, which is rarely seen in reality (Chen et al., 2021). Turbines are also installed in different regions with varying weather conditions, leading to a different statistical distribution of data that can reduce the generalizability of these models (Chen et al., 2021).

This thesis develops and evaluates a MATLAB/Simulink-based model for simulating the electromechanical systems of wind turbines and detecting faults in real time using artificial intelligence techniques. Researchers developed an integrated DFIG turbine model within Simulink to test CNN-LSTM and GAN algorithms for fault detection and compare them with traditional methods, using precise simulation settings including step size = 20 μs and the ode4 (Runge–Kutta) algorithm to analyze the dynamic performance of the system (Malhotra et al., 2016). Based on this approach, this thesis adopted the construction of a simulation model of a 12.3 kW wind turbine within the Simulink/MATLAB environment, which includes heat and vibration sensor systems linked to SVM and LSTM Autoencoder algorithms for signal analysis and operational anomaly detection, thereby achieving integration between intelligent analysis and realistic time modeling.

Mask-MRNet is a deep learning network developed to analyze blade image and accurately identify cracks (Y. Zhang et al., 2022). In some other studies, multi-channel convolutional neural networks have successfully detected blade cracks and gearbox failures with an accuracy exceeding 98% (Zare & Ayati, 2021). However, a bottleneck in the application of deep learning techniques in industrial reality is the lack of real failure data, as wind turbines typically have a long lifespan without failure(A. Wang et al., 2023). This is the essence of transfer learning facilitating the transfer of knowledge from data-rich turbines to data-poor ones, reducing the need for complete retraining of models (Dao, 2023; Rama et al., 2024). Prominent transfer algorithms such as TrAdaBoost and Inception-V3 have also proven to be highly effective in dealing with unbalanced or differentially distributed data (Ng & Lim, 2022). In another domain, automated machine learning (AutoML) techniques have been proposed to automate the process of model selection and hyperparameter tuning, which helps reduce human effort in developing diagnostic systems(Puruncajas et al., 2025; Morrison et al., 2022).

Recent studies have reported that Auto ML-ANN has the ability to achieve 96.8%–99.9% accuracy in detecting internal faults in DFIG systems, concluding that automated artificial intelligence can be implemented within the final control tuning (Machine learning-based fault detection in wind turbine generators Anonymous, 2023) (B. Wang et al., 2025; Qi et al., 2024a). Recently, modern work has focused on the use of hybrid models, which combine physical analysis and data processing methods driven by artificial intelligence algorithms to obtain realistic diagnoses(Craig et al., 2018) . For example, mechanical analysis is combined with machine learning to process SCADA data to monitor drivetrains using thermal modeling (Qin et al., 2023a). Furthermore, fuzzy reliability assessment in collaboration with machine learning has been shown to enhance failure prediction from general system condition assessments (Babu et al., 2024). These results demonstrate the potential of automated artificial intelligence technologies to improve fault detection performance and support their integration into wind turbine control and monitoring systems.

All of these academic studies converge in the spirit that AI with transfer learning and AutoML makes intelligent diagnostic systems available that can adapt to actual operating conditions and significantly reduce maintenance costs. Instead of relying on failure signals after failure, we can identify early symptoms of performance degradation by recognizing the underlying structures of SCADA data using improved deep learning models(Sheng et al., 2025). This methodology is a complete departure from proactive maintenance, which until now has been largely understood as reactive (monitoring) but rarely used as artificial intelligence (active)(Stetco et al., 2019). We conclude that the application of modern diagnostic algorithms to wind turbines using machine learning and transfer learning is not just an academic innovation but also extends to actual increases in energy efficiency, reduced emission levels, and operational sustainability (Behara & Saha, 2025a).

This study aims to create an intelligent diagnostic model that can accurately diagnose faults based on real SCADA data generated using simulink, solving both the data balance and distribution heterogeneity problems as well as learning from insufficient samples. This will help establish a scientific basis for the application of intelligent maintenance systems in modern wind farms and the future sustainable development of clean energy (García-Vaca et al., 2026).

1.2 Motivation

In recent years, the wind energy industry has experienced unprecedented expansion, with wind farms now comprising thousands of turbines operating in variable and harsh climatic conditions, increasing the likelihood of mechanical and electrical failures (Behara & Saha, 2025a; Ohunakin et al., 2024). Although traditional monitoring systems provide vast amounts of data from multiple sensors via SCADA systems, the ability to analyze this data and extract early indicators of failures remains limited, leading to significant economic losses due to unexpected downtime and high maintenance costs (Sheng et al., 2025).

The main challenge is to develop models capable of learning from historical data and predicting abnormal conditions before they occur (Stetco et al., 2019). Traditional learning algorithms such as SVM or Random Forest may achieve good results in laboratory data, but they suffer from poor performance when faced with real, unbalanced, or noisy data, as is the case with SCADA data (Chen et al., 2021). Hence, there is a need to employ more advanced techniques such as Deep Learning and Transfer Learning, which have proven their ability to extract hidden features in signals and improve diagnostic accuracy even with limited data (Zare & Ayati, 2021).

In addition, many previous studies have focused on post-failure analysis, while the practical goal is to shift towards predictive maintenance, which reduces failures before they actually occur through intelligent monitoring (Qin et al., 2023a). This trend opens the door to integrating numerical modeling in MATLAB/Simulink with artificial intelligence to build systems capable of detecting faults in real time (Laouti et al., 2011).

Based on this, the main motivation for this thesis stems from the need to develop an intelligent diagnostic system that combines physical simulation in Simulink with intelligent signal analysis using machine learning and deep learning, with the aim of increasing turbine reliability and reducing long-term operating costs.

1.3 Problem Statement

Despite significant advances in wind turbine monitoring and fault diagnosis, traditional systems still suffer from poor early fault detection and performance deviation prediction capabilities (Stetco et al., 2019). Most farms rely on SCADA data collected at infrequent intervals and subject to environmental noise and climatic

fluctuations, making it complex and difficult to analyze and extract accurate failure patterns (Sheng et al., 2025). Consequently, the main research gap is the lack of reliable models for the early detection of faults in noisy and unbalanced SCADA data (Chen et al., 2021).

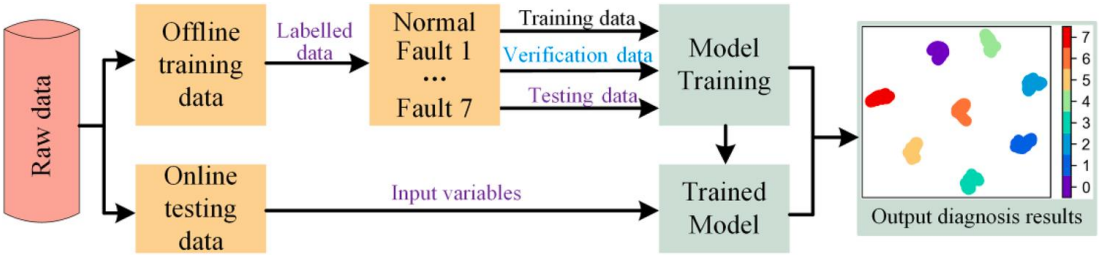


Figure 1 Flow chart of the proposed method for wind-turbine fault detection (adapted from (J. Liu et al., 2023)).

Furthermore, some studies suffer from a separation between intelligent analysis and physical modeling, where analyses are performed using ready-made data without dynamic representation of the system, which reduces the reliability of the results when applied to a real operating environment (Craig et al., 2018). In contrast, building a dynamic simulation model in the MATLAB/Simulink environment allows for the generation of realistic data that mimics the behavior of the actual turbine, enabling AI algorithms to be tested on a larger scale and with greater accuracy (Laouti et al., 2011).

On the other hand, the problem of model generalization remains, as algorithms perform well on specific data but fail when transferred to different turbines or new environments (Ng & Lim, 2022). Therefore, there is a need to integrate technologies such as transfer learning and deep learning into diagnostic systems to improve their ability to learn from limited data and adapt to operational differences (A. Wang et al., 2023).

Based on this, the problem lies in developing an integrated intelligent diagnostic model based on dynamic simulation in Simulink/MATLAB and using advanced learning algorithms to analyze vibration and heat signals, in order to achieve early and reliable detection of faults in wind turbines and reduce losses resulting from unplanned downtime (Al Lahham et al., 2025).

1.4 Aims

This research aims to develop an integrated intelligent diagnostic model for wind turbines based on combining dynamic simulation in the MATLAB/Simulink environment with artificial intelligence algorithms for early and accurate detection of mechanical and electrical faults. The model is designed to address the issues identified in previous studies related to data imbalance, difficulty in generalization, and poor performance of traditional models in dealing with real-world SCADA data (Sheng et al., 2025). The thesis also aims to build a simulation database that represents the actual physical behavior of the turbine through heat and vibration signals generated by the dynamic system simulated in Simulink, allowing the SVM and LSTM Autoencoder algorithms to be tested under varying operating conditions to measure their efficiency in detecting early operational deviations.

The detailed objectives of the study are as follows:

Create an integrated simulation model of a 12.3 kW wind turbine using MATLAB/Simulink to generate dynamic data that simulates operational reality.

Apply machine learning and deep learning algorithms (Autoencoder, SVM, ANN, and CNN-LSTM) to analyze heat and vibration data and diagnose faults in their early stages.

Integrating the results of intelligent analysis with the simulation environment to develop an integrated predictive maintenance framework that contributes to reducing unexpected downtime and improving the operational efficiency of wind energy systems.

Evaluating the performance of the proposed models using accuracy, sensitivity, and error rate metrics, and comparing them with traditional methods.

1.5 Thesis Contributions

This thesis presents a scientific and practical contribution to the field of intelligent fault diagnosis for wind turbines through the development of an integrated simulation model based on the MATLAB/Simulink environment and artificial intelligence algorithms. The work focuses on designing an early diagnosis system that combines physical modeling and data-driven analysis, with the aim of detecting operational deviations before they turn into actual faults (Qin et al., 2023a).

The first contribution is the construction of a dynamic simulation model of a 12.3 kW wind turbine capable of generating data that mimics the actual behavior of the turbine under different operating conditions, enabling the accurate and safe study of mechanical and electrical failures (Dallabona et al., 2025).

The second contribution is the application of hybrid machine learning and deep learning algorithms such as SVM and LSTM Autoencoder to analyze vibration and heat signals emitted by the turbine, and use them to design an intelligent system for early fault detection and improved real-time diagnostic accuracy (Zare & Ayati, 2021).

The third contribution is the development of a Predictive Maintenance Framework that integrates the results of intelligent analysis with simulation data, enabling the prediction of fault indicators before they occur and proactive maintenance decisions that reduce downtime and increase operational reliability (Babu et al., 2024).

Finally, the thesis represents a step toward the transition from passive monitoring to intelligent self-diagnosis, through a model that is potentially applicable in future industrial environments for wind energy systems.

1.6 Thesis Outlines

This thesis is logically organized into six consecutive chapters, providing a comprehensive and sequential study on the intelligent diagnosis of wind turbine faults using the MATLAB/Simulink environment and artificial intelligence techniques.

Chapter II: provides a comprehensive review of the previous literature, highlighting traditional and modern methods for detecting faults in wind turbines, with a focus on SCADA data-based monitoring systems, vibration and heat analysis, and machine learning methods used in operational condition assessment.

Chapter III: presents the methodology and system design adopted in this research, explaining the conceptual framework, modeling mechanism, and structure of the hybrid diagnostic system that combines physical simulation and intelligent data analysis.

Chapter IV: describes the simulation process and model implementation in the MATLAB/Simulink environment, explaining the dynamic model of a 12.3 kW wind turbine, sensor integration, system parameter setup, and the introduction of fault scenarios to generate realistic operating data for subsequent analysis.

Chapter V: presents the results of the analysis using AI-based diagnostics, explaining how to process simulation data through SVM and LSTM Autoencoder algorithms, evaluate performance, and compare the proposed model with traditional methods to verify its accuracy and efficiency.

Chapter VI: concludes the thesis by presenting the most important results and general conclusions, highlighting the main contributions, recommendations, and future directions for research toward more intelligent and predictive maintenance systems in wind energy applications.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter aims to review and analyze previous studies related to fault diagnosis in wind turbines and to review recent trends in integrating numerical modeling with artificial intelligence to improve the reliability of renewable energy systems. This chapter focuses on the key concepts that support this research, starting with the components of the turbine and how it works, moving on to traditional methods of monitoring energy systems, and ending with developments in the fields of machine learning and deep learning that have made predictive maintenance possible and effective. The chapter also compares experimental and numerical models and explains the importance of the MATLAB/Simulink environment in studying the dynamic performance of turbines and analyzing signals resulting from vibration and heat. It also highlights the evolution from manual diagnostic methods to intelligent systems capable of learning from data, representing a shift from passive monitoring to proactive diagnostics. This chapter establishes the scientific basis for understanding the current research problem and identifies gaps in the previous literature that this work seeks to fill through an integrated simulation model based on artificial intelligence and simultaneous physical analysis.

2.2 Overview of the Wind Turbine System

A wind turbine system consists of a number of mechanically and electrically interconnected parts that work together to convert the kinetic energy generated by the wind into usable electrical energy (Qi et al., 2024a; Xiang et al., 2021).

The main components include:

Rotor Blades: Capture wind energy and convert it into rotational motion (Binti Shahrulhisham et al., 2022).

Gearbox: Transfers mechanical torque and increases rotational speed to match the generator (Zare & Ayati, 2021).

Generator: Converts mechanical energy into electrical energy (Yin et al., 2020).

Sensors and Control Unit: Monitors vibration, temperature, and wind speed to control operational stability (An et al., 2011; Ghanbari et al., 2023).

The mechanical energy extracted from the wind can be expressed by the basic physical relationship:

$$P = \frac{1}{2}\rho A v^3 C_p \quad (1)$$

The main bearing is one of the most important mechanical components in the drivetrain system, as it bears the weight of the rotor and the torque generated by the wind, and works to stabilize and stabilize the rotational movement. Any malfunction in this part leads to irregular vibrations and loss of rotor balance, which affects the turbine's performance and operational efficiency (Ghanbari et al., 2023).

The performance of the main bearing is affected by several factors, including:

Air temperature (T_{air}) and coolant temperature (T_{liq}).

Mechanical friction (Q friction) resulting from torque and rotor speed.

Heat transfer coefficient (h) for both oil and air.

These factors were studied mathematically using a partial differential equation (PDE) to characterize heat transfer within the bearing as follows (Malik, 2024) :

$$\frac{\partial T(t,x,y)}{\partial t} = \alpha\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right) - \frac{h_{air}}{C_p}(T(t,x,y) - T_{air}) - \frac{h_{liq}}{C_p}\left(T(t,x,y) - T_{liq}\right) + \frac{Q_{friction}}{C_p} \quad (2)$$

This equation explains that the change in bearing temperature depends on thermal diffusion, air and liquid cooling, and heat generated by friction.

2.2.1 Traditional Monitoring Methods

Turbine monitoring previously relied on threshold-based alarm systems and periodic manual inspections (Craig et al., 2018; An et al., 2011). SCADA systems record operating variables such as temperature and rotor speed at intervals of typically ten minutes, and safe operating limits are determined based on these readings (Chen et al., 2021). However, these methods often fail to detect slow-developing faults, as

low-resolution data and environmental noise mask subtle changes (Sheng et al., 2025). Vibration analysis is one of the most reliable traditional methods for detecting mechanical faults in bearings and gearboxes (Ghanbari et al., 2023). Acceleration sensors are mounted on the main bearing and gearbox, and the signals are analyzed in the frequency domain to detect imbalance or misalignment (B. Wang et al., 2025). However, vibration data is high-dimensional and requires advanced signal processing to distinguish between normal and abnormal conditions (X. Zhang et al., 2020). Temperature monitoring is another traditional technique that uses thermocouples near bearings and generator coils to detect excessive heat generated by friction or electrical losses.(B. Wang et al., 2025). Used a heat transfer model to analyze the thermal distribution within the main bearing, showing that temperature variations are an early indicator of failures. Partial differential equations (PDEs) are often used to represent heat diffusion and cooling effects within the bearing (Malik, 2024).

2.2.2 SCADA-based monitoring

With the proliferation of SCADA (Supervisory Control and Data Acquisition) systems, wind farms have begun to collect vast amounts of operating data such as rotor speed, pitch angle, output power, ambient temperature, and wind speed (Bindingsbø et al., 2023; Stetco et al., 2019). SCADA-based condition monitoring is characterized by its low cost and the availability of data over long periods (Pozo & Vidal, 2015). However, its low sampling rate (typically every 10 minutes) limits its accuracy in detecting rapid mechanical failures such as bearing wear or gear tooth cracking (Sheng et al., 2025). To overcome this limitation, researchers are combining SCADA data with high-frequency vibration and temperature signals to develop more comprehensive diagnostic models (Ghanbari et al., 2023).

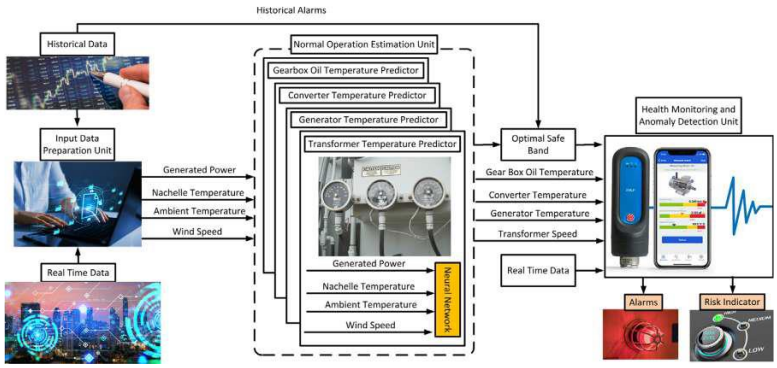


Figure 2 Temperature-based condition monitoring framework for wind turbines ((Al Lahham et al., 2025)

2.2.3 Data-Driven Monitoring

In recent years, artificial intelligence (AI) methods have replaced traditional methodologies (Stetco et al., 2019). Machine learning techniques rely on extracting patterns from historical data to distinguish healthy from damaged conditions without the need for a specific physical model (Pozo & Vidal, 2015). The most prominent of these methods are: Support Vector Machines (SVM), Decision Trees, and Artificial Neural Networks (ANN) (Chen et al., 2021). Recent studies have also expanded to use deep learning, such as long short-term memory (LSTM) and autoencoders, which are capable of analyzing temporal changes in temperature and vibration data with high accuracy(Yan et al., 2024; C. Zhang & Yang, 2023). These models are characterized by their ability to handle nonlinear relationships between variables and identify hidden faults within raw signals (Rudin et al., 2011).

2.2.4 Conclusion

Condition monitoring technologies in wind turbines have evolved from simple alarm-based systems to integrated intelligent systems that combine vibration, temperature, and SCADA data using artificial intelligence algorithms (Stetco et al., 2019). Although traditional methods remain useful for initial diagnosis, they suffer from low accuracy and time delays (Sheng et al., 2025). Modern AI-based methods enable more effective and reliable predictive maintenance in large wind farms (Behara & Saha, 2025a).

2.3 SCADA-Based Fault Detection

2.3.1 Introduction to SCADA in Wind Turbines

The Supervisory Control and Data Acquisition (SCADA) system is the cornerstone of modern wind turbine monitoring and operation (Stetco et al., 2019). This system collects operational data from hundreds of sensors distributed throughout the turbine in the gearbox, main bearing, generator, and pitch system and converts it into digital signals that are sent to a central control unit for real-time performance monitoring (Pozo & Vidal, 2015). The measured data is stored periodically at regular intervals (usually every 10 minutes) and includes variables such as wind speed, rotor torque, pitch angle, output power, oil temperature, bearing and generator temperature, and others (Chen et al., 2021). This vast store of data is used to analyze the operating

condition and evaluate the thermal and mechanical efficiency of the turbine (Sheng et al., 2025).

The SCADA system is characterized by its ability to remotely monitor faults and send immediate notifications when specified values are exceeded, reducing the need for on-site inspection and contributing to lower operating costs (Pozo & Vidal, 2015). Historical data analysis also allows current performance to be compared with normal reference values, helping to detect early deviations that may indicate mechanical or electrical malfunctions (Stetco et al., 2019).

The SCADA architecture typically consists of three main levels:

Sensors & Transducers, which collect raw information.

The field controller (PLC), which receives signals and converts them into digital data.

The control center, which stores data, displays it on the user interface, and analyzes it using supervisory and control algorithms.

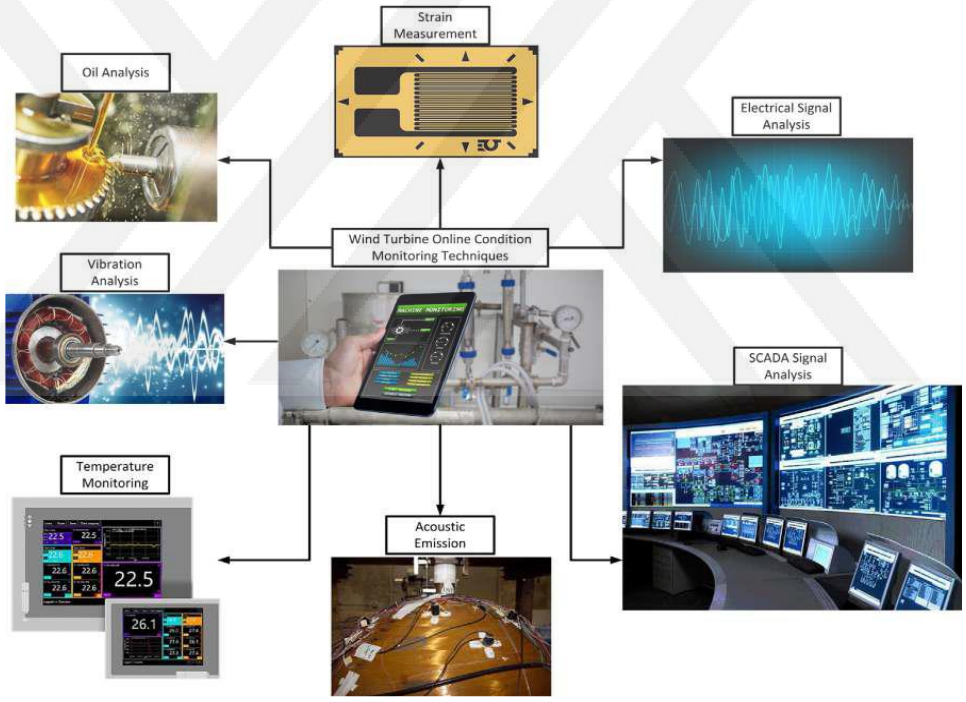


Figure 3 SCADA monitoring techniques in wind turbines ((Al Lahham et al., 2025)

This figure 3 illustrates the general structure of the SCADA system, showing the layers of data collection and transmission to the control center, and the relationship between operational variables (speed, torque, temperature) and control units. The SCADA system provides a solid foundation for the development of smart monitoring

systems, as the thermal and mechanical data it generates can be integrated with artificial intelligence algorithms to build predictive diagnostic models that go beyond the limitations of traditional alarm systems (Stetco et al., 2019). This feature makes SCADA an essential component in predictive maintenance research, as it is the layer that feeds machine learning algorithms with highly relevant raw data (Ng & Lim, 2022).

2.3.2 Fault detection mechanism using SCADA data

SCADA systems in wind turbines play a pivotal role in the early detection of faults by analyzing operational data that is continuously collected from turbine components (Chokr et al., 2024; Sheng et al., 2025). The mechanism typically begins with the measurement and data collection phase, where key variables such as generator and bearing temperature, wind speed, current, voltage, and torque are recorded (Chen et al., 2021). This data is then transferred to a central database for analysis using statistical algorithms or advanced machine learning techniques (Pozo & Vidal, 2015).

The principle of fault detection is based on comparing real-time measured values with reference models or historical operating patterns recorded during normal operation (Stetco et al., 2019). Any clear deviation between actual and expected behavior indicates an abnormal condition that warrants close monitoring or preventive shutdown (Sheng et al., 2025). Various indicators called fault indicators are used in this process, such as rates of change in temperature or the correlation coefficient between vibration and power output (Ghanbari et al., 2023).

$$D^2 = (\boldsymbol{x} - \boldsymbol{\mu})^T \, \boldsymbol{\Sigma^{-1}} \, (\boldsymbol{x} - \boldsymbol{\mu}) \tag{3}$$

This is the Mahalanobis Distance equation, which is used to determine the extent to which the current state deviates from the healthy reference state (J.-H. Liu et al., 2022; Rudin et al., 2011). Methods based on multivariate statistical analysis, such as Principal Component Analysis (PCA) and Independent Component Analysis (ICA), are used in the analysis phase (Pozo & Vidal, 2015). These methods are applied to SCADA data to reduce dimensions and extract the variables most sensitive to faults (Rossi et al., 2016). This step is necessary to identify influential variables before training any predictive or diagnostic model (Stetco et al., 2019). After extracting the

characteristic indicators, classification or prediction algorithms are used to determine the status:

- Normal Operation
- Warning
- Fault

These states are determined based on thresholds defined from historical operating data (Rossi et al., 2016). This mechanism makes the SCADA system an effective tool for predictive maintenance, as it can identify faults before actual mechanical failure occurs, thereby avoiding turbine downtime and reducing energy losses (Behara & Saha, 2025a).

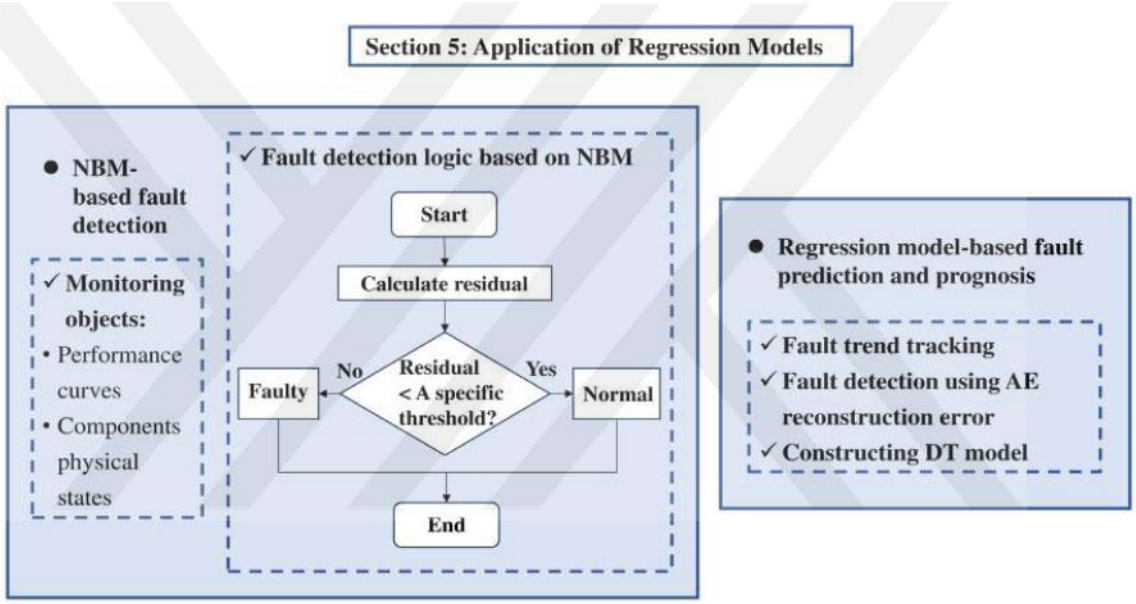


Figure 4 Fault detection mechanism using SCADA data (regression decision logic) ((Sheng et al., 2025), Neurocomputing, 633, 129830, Fig. 10)

The figure illustrates the fault detection logic using SCADA data, where the process begins by calculating the residuals between the actual and expected values. If the residual value exceeds a certain threshold, the condition is classified as “Faulty”; otherwise, it is considered “Normal” (Sheng et al., 2025).

2.3.3 Challenges and limitations in SCADA systems

Despite the significant importance of SCADA systems in monitoring wind turbines and diagnosing their malfunctions, their use in research or industrial applications faces a number of technical and practical limitations that affect the accuracy and reliability of the results (Stetco et al., 2019). The first of these challenges is the low sampling rate. SCADA data is typically recorded at intervals ranging from 1 to 10 minutes, making it unable to capture rapid changes in signals such as sudden vibrations or momentary temperature spikes, which leads to the loss of critical details that could aid in early fault detection (Sheng et al., 2025).

The second challenge is the presence of noise and measurement errors in the data, as SCADA signals are affected by environmental factors such as humidity, temperature, and network voltage fluctuations (Pozo & Vidal, 2015). These disturbances lead to unstable or contradictory data, requiring complex cleaning and statistical analysis before they can be used in diagnostic models (Rossi et al., 2016). Data gaps are also a common problem in SCADA systems, resulting from communication interruptions or sensor failures. These gaps reduce the consistency of time series and make it difficult to accurately train intelligent models (J. Liu et al., 2023). From an organizational perspective, SCADA systems suffer from data confidentiality constraints, as energy companies consider sensitive operational data to be proprietary and cannot be shared for research or analytical purposes, which limits the availability of open-source datasets that can be used to develop new algorithms (Ng & Lim, 2022).

Finally, there is a challenge related to the integration between SCADA systems and other smart systems (Integration Challenges). Older SCADA systems often operate on closed or limited communication protocols, making their integration with modern AI-based analysis systems complex in terms of software and structural compatibility (Craig et al., 2018). Overall, these challenges are one of the main reasons why researchers rely on modeling and simulation in MATLAB/Simulink as a digital alternative for generating clean, structured virtual data used in the development of smart diagnostic systems (Al Lahham et al., 2025).

2.3.4 Recent developments in integrating SCADA systems with artificial intelligence

Recent years have seen a growing trend toward integrating SCADA systems with artificial intelligence (AI) technologies to develop more efficient predictive diagnostic and maintenance systems for wind turbines (Stetco et al., 2019). While SCADA's role was previously limited to collecting operational data and analyzing it statistically, today it is used to train intelligent models that automatically learn patterns of proper performance and recognize anomalies before they turn into actual failures (Ng & Lim, 2022).

Most recent studies rely on converting SCADA data into machine learning models that use algorithms such as Support Vector Machine (SVM), Artificial Neural Networks (ANN), and Extreme Gradient Boosting (XGBoost) to classify operational states as “normal” or “abnormal,” or to predict future performance indicators based on historical operational data (Chen et al., 2021; Fadzail et al., 2025).

Among the most notable developments is the use of Deep Neural Networks, particularly Long Short-Term Memory (LSTM) algorithms, which are characterized by their ability to handle long-term time series data such as that generated by SCADA (C. Zhang & Yang, 2023). LSTM is used to predict future temperatures or vibrations based on data from previous days, thus enabling the early detection of abnormal thermal or mechanical trends (Reddy et al., 2024).

In other recent studies, SCADA data has been combined with data from other monitoring systems such as CMS (Condition Monitoring System) to form what are known as Hybrid Data-Driven Models (Al Lahham et al., 2025). These models are used to integrate multiple signals (temperature, vibration, pressure, speed) within a single framework for analysis, improving detection accuracy to over 95% in some experiments (Allal et al., 2024).

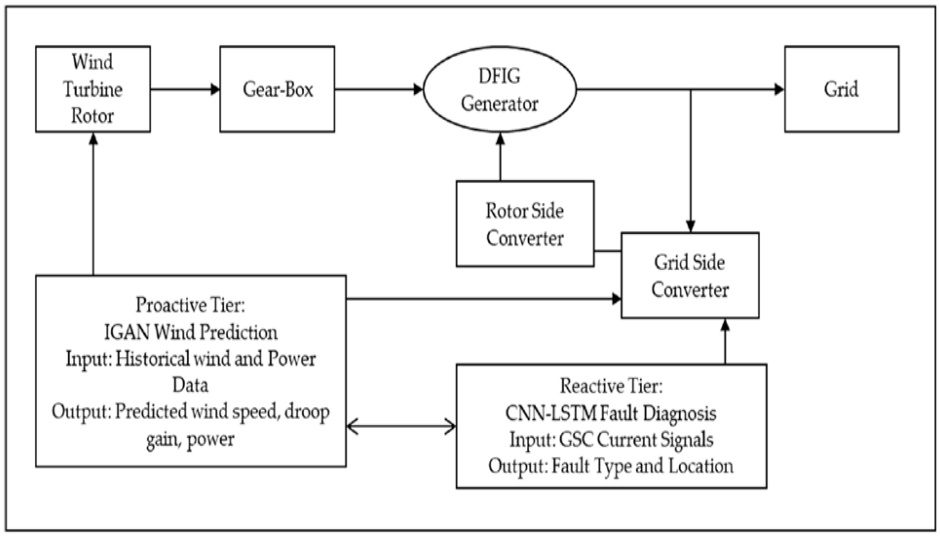


Figure 5 Two-layer framework for integrating SCADA data with artificial intelligence in wind turbine systems (Behara & Saha, 2025a)

The figure illustrates the general architecture of a dual-layer intelligent system that combines the IGAN model for proactive wind and energy prediction with the CNN–LSTM model for real-time fault detection using SCADA data, enabling improved operational stability and accurate prediction of the future state of the turbine (Behara & Saha, 2025a).

The use of unsupervised learning algorithms such as K-Means Clustering and Autoencoders to collect data and detect anomalies without the need for prior classifications has also emerged, which is an important step towards the application of self-monitoring in large wind farms (Chen et al., 2021).

Through these developments, AI-powered SCADA systems are now capable of:

- Analyzing large amounts of data in real time.
- Predicting failures with high temporal accuracy.
- Supporting automated decision-making in operation and maintenance.

However, the effectiveness of these systems still depends mainly on the quality of the data input from SCADA, which makes combining modeling and simulation (as in MATLAB/Simulink) with actual operational data one of the most important current research trends in the field of renewable energy (Al Lahham et al., 2025).

2.3.5 Conclusion

A review of recent literature shows that wind turbine fault monitoring and diagnosis has undergone significant development over the last decade, particularly following the integration of artificial intelligence technologies with SCADA system data. While traditional monitoring relied on analyzing limited signals such as temperature and vibration within a narrow time frame, today it includes advanced machine learning systems capable of analyzing massive amounts of real-time data and extracting accurate operating patterns.

There has also been a trend toward developing hybrid models that combine multiple data sources (temperature, vibration, pressure, wind speed) within an integrated learning framework, which has improved diagnostic reliability and predictive maintenance efficiency by more than 90% in several studies.

Despite these developments, challenges remain in terms of data reliability and quality, as well as the need to expand the integration between dynamic simulation systems such as MATLAB/Simulink and actual operational data to create more realistic digital environments for predicting future failures.

Overall, the results confirm that combining SCADA and artificial intelligence is one of the most promising research trends toward creating self-monitoring, self-learning systems capable of reducing downtime and improving the operational efficiency of wind turbines.

CHAPTER THREE

METHODOLOGY AND SYSTEM DESIGN

3.1 Introduction

This chapter aims to present the proposed research methodology for diagnosing faults in wind turbines using artificial intelligence techniques, within an integrated simulation environment developed in MATLAB/Simulink. The work focuses on building a realistic numerical model of a 12.3 kW wind turbine that simulates the electromechanical and thermal aspects of the system, including generator behavior, component vibration, and temperature rise due to load and friction. This model aims to generate operational data similar to that of industrial SCADA systems, providing a reliable research alternative for training and testing AI models under varying operating conditions and multiple faults.

The proposed methodology is based on a Multi-Layer AI Framework that integrates deep learning and traditional machine learning techniques. In the first stage, LSTM and CNN algorithms are used to analyze the temporal and mechanical signals extracted from the turbine and identify critical patterns that precede failures. The results of the models are then integrated into a multi-layer neural network (ANN) to generate an accurate classification between normal and abnormal operation, while an SVM algorithm is used to verify the results and improve decision boundaries. Finally, an Autoencoder model is employed to detect previously unclassified anomalies, enhancing the system's ability for intelligent self-monitoring.

This proposed system contributes to the construction of a predictive diagnostic framework capable of early fault detection, reducing downtime, and improving turbine reliability. The following sections provide details on model design, algorithm training, and results analysis, which will be discussed in detail in Chapter 4.

3.2 Overall Research Methodology

This section presents the overall methodological framework followed in this research to develop an intelligent system for diagnosing faults in wind turbines. The methodology was designed to cover the physical, engineering, and intelligent aspects of the system, starting with the construction of the mathematical model in the

MATLAB/Simulink environment, through the collection of virtual operational data, and ending with the application and analysis of various artificial intelligence algorithms.

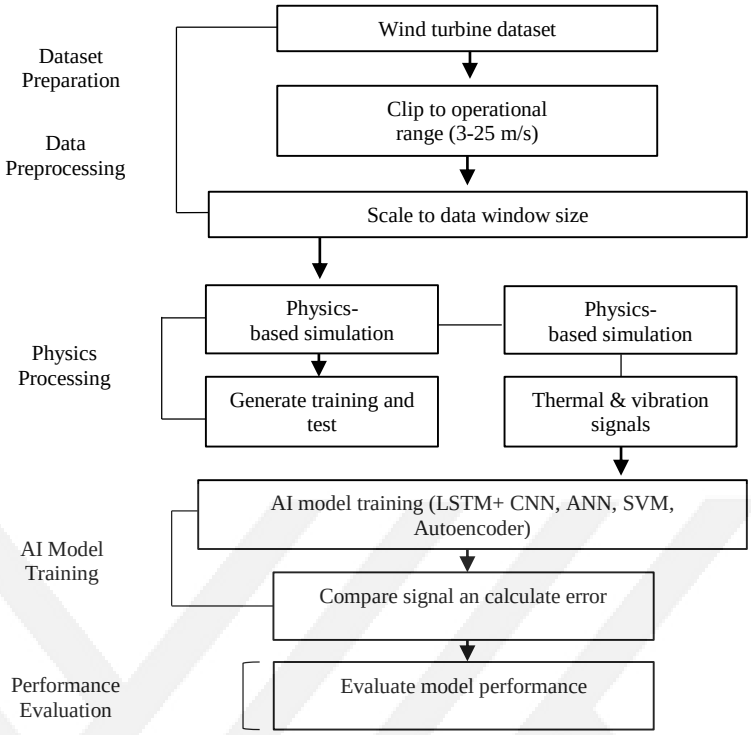


Figure 6 Overall methodology steps

Turbine model design in MATLAB/Simulink: A dynamic model of a 12.3 kW turbine was built, including mechanical, electrical, and thermal systems, with the integration of a heat and vibration sensor unit and a blade angle control motor. The model allows monitoring of changes resulting from hypothetical failures.

Generation of operational data: Simulation is used to generate realistic time data including rotational speed, power output, temperatures, and vibrations. This data represents an operating environment similar to SCADA data and is used to train intelligent models (B. Wang et al., 2025).

Signal processing and feature extraction: Statistical and temporal transformations are applied to extract features such as RMS, STD, Mean, Energy, and Max, which represent the behavior of the system in different states.

Application of Artificial Intelligence Algorithms: Several models were trained, including LSTM, CNN, ANN, and SVM, in addition to Autoencoder to detect

abnormal patterns. Hybrid learning was performed to improve the overall performance of the system (Sheng et al., 2025; B. Wang et al., 2025).

Verification and analysis: The results of the models are compared using performance metrics such as Accuracy, RMSE, and Loss Function, with Confusion Matrices analyzed to evaluate the efficiency of each algorithm.

These steps combine a simulation-based physical approach with a data-driven intelligent approach, enabling the proposed system to operate in a real or virtual environment with high efficiency.

3.3 Wind Turbine System Modeling in Simulink

This section aims to explain the dynamic and electromechanical structure of a wind turbine system designed in the MATLAB/Simulink environment.

The system consists of a series of interconnected physical blocks that represent the generation of mechanical power from wind and its conversion into electrical energy that can be connected to the grid. The model was implemented to simulate the behavior of a 12.3 kW turbine under normal operating conditions and hypothetical fault conditions.

3.3.1 Wind Turbine System Modeling in Simulink

The aerodynamic model is based on the equation for power generated from wind:

$$P_{wt} = \frac{1}{2}\rho A v^3 C_p(\lambda, \beta) \quad (4)$$

where ρ represents air density, A represents rotor area, v represents wind speed, and Cp is the performance coefficient that varies according to the tip speed ratio:

$$\lambda = \frac{\omega_r R}{v} \quad (5)$$

The aerodynamic power is converted into mechanical torque $T_m = P_w t/\omega_r$ that feeds the generator shaft.

The mechanical dynamics of the system represent the relationship between torque and speed using the Lagrange equation:

$$J\frac{d\omega_r}{dt} = T_m - T_e - B\omega_r \quad (6)$$

Where M, D, and K are the mass, damping, and stiffness matrices, respectively, and q(t) is the overall displacement of the blades (Laouti et al., 2011).

The main simulation parameters used in the proposed wind turbine model are summarized in Table 1. These parameters were selected based on commonly reported wind turbine operating conditions in the literature and adjusted empirically to ensure stable and physically consistent simulation behavior under varying operating and fault conditions.

Table 1. Wind Turbine Simulation and Training Parameters.

Parameter	Value
Rated turbine power	12.3 kW
Air density	1.225 kg/m³
Wind speed operating range	3–25 m/s
Solver type	Fixed-step
Simulation approach	Discrete-time
Sampling interval	1 day
Sliding window length	64
CNN filters	32
Kernel size	3
LSTM hidden units	64
Batch size	128
Learning rate	0.001
Training epochs	200

3.3.2 Generator and Electrical Subsystem

The generator is represented as a permanent magnet synchronous generator (PMSG) connected via a three-phase rectifier bridge and then a boost DC–DC converter (Villanueva & Feijóo, 2020).

$$\begin{cases} v_d = R_s i_d - \omega_e L_q i_q + L_d \frac{di_d}{dt} \\ v_q = R_s i_q + \omega_e L_d i_d + L_q \frac{di_q}{dt} + \omega_e \psi_f \end{cases} \quad (7)$$

$$T_e = \frac{3}{2} p\left(\psi_f i_q + \left(L_d - L_q\right) i_d i_q\right) \quad (8)$$

The converter is controlled by a Perturb & Observe MPPT algorithm to adjust the duty ratio in order to maximize the output power.

The electrical system consists of:

- Three-phase voltage and current measurement
- Universal Bridge unit for current conversion
- DC-Link Capacitor for output smoothing

$$C_{dc} \frac{dV_{dc}}{dt} = i_{in} - i_{load}, \qquad V_{out} = \frac{V_{in}}{1-D} \quad (9)$$

- Boost Converter for voltage boosting before feeding the load.

3.3.3 Control Strategy

A controller based on Vector Control (dq Transformation) was applied to convert signals from the fixed frame to the rotating reference, allowing for high-precision control of torque and power transfer.

Speed and torque values are adjusted within optimal limits using PI Controllers with protection limiters in case of voltage or current increase (Eriksson et al., 2008).

3.3.4 Integration with Thermal and Vibration Systems

The model also includes thermal units (Thermal Mass + Heat Flow Source) to represent heat rise within the generator, and a mechanical vibration unit that simulates the system response to imbalance or weak load (B. Wang et al., 2025). Both the heat and vibration signals are sent to the smart sensor unit described in the next section.

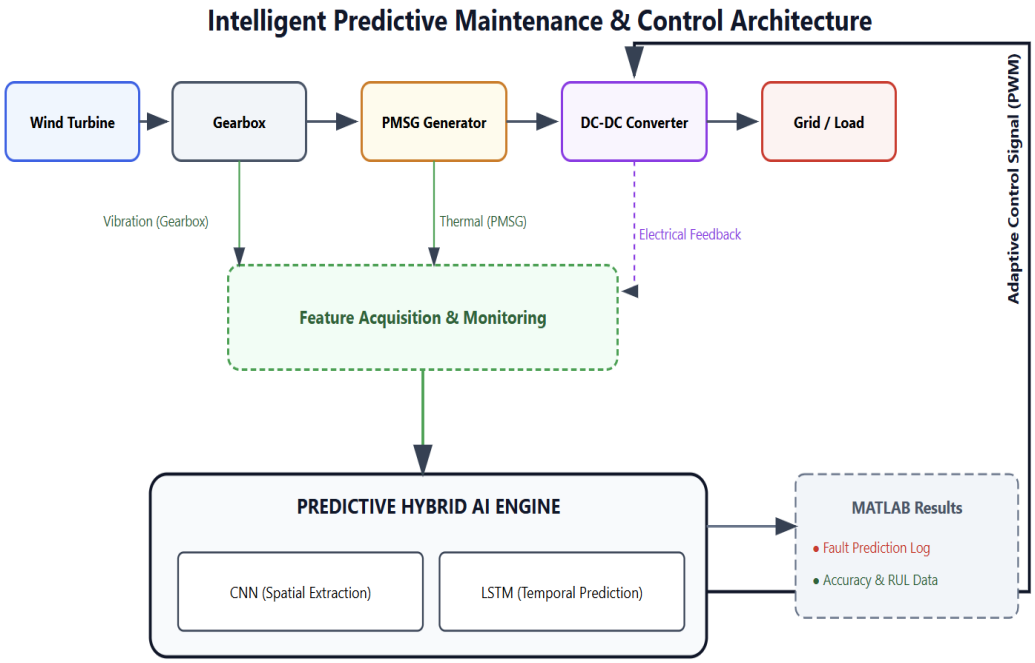


Figure 7 Complete physical model of a wind turbine in the Simulink environment, showing the turbine, generator, power converter, and control and measurement system units.

3.4 Sensing and Monitoring Subsystem

This part aims to build a measurement and monitoring layer based on heat and vibration sensors, with a data logging system that feeds the artificial intelligence algorithms with the variables needed for early fault detection. The architecture was implemented within the MATLAB/Simulink environment according to a scheme that ensures integration with the electromechanical model of the turbine (Laouti et al., 2011).

3.4.1 Temperature Sensing

a) **Measurement location:** The Temperature Sensor was connected to the generator's Thermal Mass path via the PS-Simulink Converter and then to the signal conditioning path (Gain/Scope) as shown in the diagram. This architecture represents the internal heating resulting from generator losses and heat exchange with the surrounding environment.

b) **Measured signal:** Temperature T(t) in °C, recorded via Scope11 and output to MATLAB using To Workspace for training and analysis and the data was collected under normal operating conditions.

c) **Measurement objective:** Track slow thermal drift resulting from harsh operating conditions or the onset of bearing degradation, with alarm thresholds set in the artificial intelligence layer.

3.4.2 Vibration Sensing

a) **Measurement location:** The signal is captured from the Ideal Rotational Motion Sensor on the generator shaft.
b) **Acceleration extraction:** The rotational output is converted to equivalent linear acceleration using Discrete Derivative and a standard gradient (Gain), then passed through adaptation stages (Abs/Filter) before being sent for processing (Ghanbari et al., 2023).

Measured signal: Vibration acceleration a(t) expressed in m/s^2. The rotational motion of the turbine components was converted to equivalent linear acceleration signals based on vibration sensor measurements. The signals were displayed using the Scopes tool and exported to MATLAB via the "To Workspace" block, while simultaneously being sent to the artificial intelligence (CNN/LSTM) processing path within the model.

3.4.3 Auxiliary Signals

To improve the distinguishability between cases, the electrical power P(t), DC link voltage Vdc(t), and mechanical speed ωr(t) are also recorded from the measurement blocks (Three-Phase VI Measurement, DC-Link, Scope) for later use in feature extraction or as additional input variables for the models.

3.4.4 Wind Input & Conditioning

Source: Real wind series for Istanbul for 2024 (365 days), imported into MATLAB and passed to the model via the From Workspace block.

Conditioning: The series was clipped to the turbine operating range [3,25] m/s and the clipping and input limits were respected to ensure realistic simulation.

Connection: The signal feeds the 12.3 kW Wind Turbine block as a wind speed controller, with a branch path for statistical processing and threshold comparison.

3.4.5 Fault Injection Scenarios

For training and testing purposes, four operational scenarios were defined:

- **Normal:** Operation within thermal and vibration limits.
- **Fault-Heat (Drift):** Gradual thermal increase at the Thermal Mass block leading to T(t) rising above a specified safe limit.
- **Fault-Vibration (Impulsive/Burst):** Short, high-demand vibration pulses on the acceleration line a(t) representing imbalance or bearing damage initiation.
- **Fault-Combined:** Thermal drift coinciding with increased vibration; the most difficult cases to classify.

Each time window is tagged with a binary/multiclass classification within the preparation path (MATLAB Function → toLabel block) before being fed into the models.

3.4.6 Data Logging and Management

A combination of Scopes + To Workspace is used to collect the series T(t), a(t), P(t), Vdc(t), ωr(t) in time-stamped MAT files. A direct path is also used to feed the networks (LSTM/CNN) inside Simulink via Buffer/Subsystem blocks for quasi-real-time testing. Training, verification, and testing copies are saved with a standard split (e.g., 70/15/15) to ensure no data leakage.

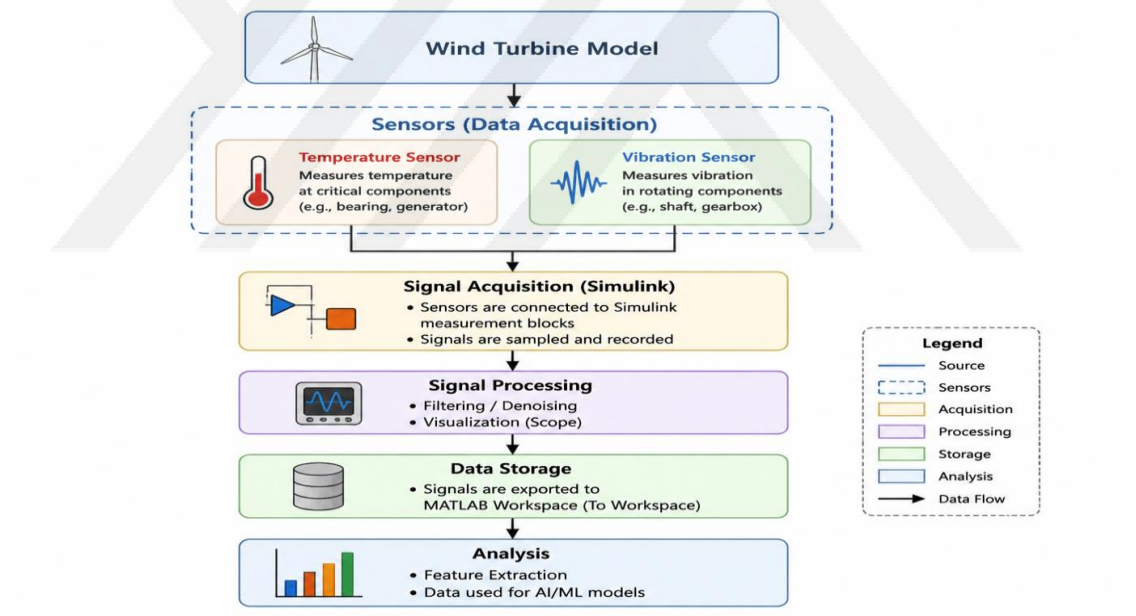


Figure 8 Data acquisition and processing workflow for temperature and vibration signals in the wind.

3.5 Proposed AI Model

This section aims to explain the proposed intelligent mechanism for fault detection in a wind turbine system designed in the MATLAB/Simulink environment.

After implementing the physical mechanical and electrical model and collecting temperature and vibration data from thermal and vibration sensors, a machine learning system was developed that is capable of analyzing time data and detecting abnormal patterns with high accuracy.

The proposed methodology is based on the use of a set of advanced artificial intelligence algorithms, including (Sheng et al., 2025; *Wang et al., 2025*, n.d.):

- **Long Short-Term Memory (LSTM)**
- **Convolutional Neural Network (CNN)**
- **Artificial Neural Network (ANN)**
- **Support Vector Machine (SVM)**

And **Autoencoder**, with the aim of comparing their performance in predicting anomalies and detecting faults before they occur.

The models were trained using data extracted from a simulation environment covering 365 days of actual wind speeds in Istanbul for the year 2024, with heat and vibration signals generated by the physical model recorded in multiple operating conditions (normal, gradual thermal failure, and instantaneous vibration failure).

The data was then divided into training and testing sets at a ratio of 70% / 30% to ensure generalization accuracy and achieve high reliability in the diagnostic results.

3.5.1 Long Short-Term Memory (LSTM)

The Long Short-Term Memory (LSTM) algorithm is one of the most important models of Recurrent Neural Networks (RNN) and is used with high efficiency in processing time-series data such as temperature and vibration data extracted from wind turbines. LSTM is characterized by its ability to learn long-term relationships between data points, making it ideal for detecting gradual failures or abnormal trends in operating signals (B. Wang et al., 2025).

The LSTM structure is based on a series of cells that contain three main components for controlling the flow of information:

- **Input Gate:** Controls the amount of new information added to the cell state.
- **Forget Gate:** Determines whether old information should be retained or discarded.
- **Output Gate:** Generates the current output based on the updated cell state.

The learning process within the LSTM unit is expressed mathematically according to the following equations:

$$\begin{cases} f_t = \sigma\left(W_f[h_{t-1}, x_t] + b_f\right) \\ i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \\ \widetilde{c_t} = \tanh(W_c[h_{t-1}, x_t] + b_c) \\ c_t = f_t \odot c_{t-1} + i_t \odot \widetilde{c_t} \\ o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \\ h_t = o_t \odot \tanh(c_t) \end{cases} \quad (10)$$

Where: ft: Forget gate, it: Input gate, ot: Output gate, ct: Cell State, ht: Cell output at time t, σ: Sigmoid Function, tanh: Hyperbolic Tangent Function

In this research, the LSTM algorithm was used to analyze the time signals generated by the temperature and vibration sensors in order to predict abnormal conditions that may indicate the onset of malfunctions.

The model was trained using 365 days of data divided into short time segments (Windows) to learn the normal behavior of the system and compare it with abnormal conditions.

A Mean Squared Error (MSE) loss function was also used during training, with the Adam Optimizer algorithm to gradually adjust the weights(B. Wang et al., 2025) .

Figure (3.4): Shows the internal structure of the LSTM algorithm used to analyze heat and vibration signals to predict abnormal conditions in wind turbines.

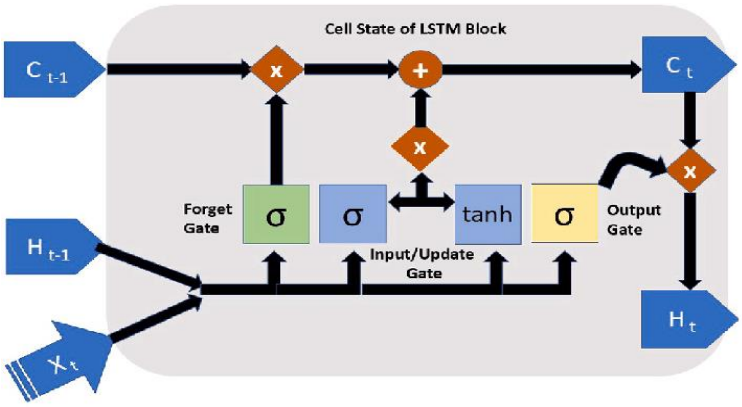


Figure 9 Internal structure of the LSTM algorithm used in analyzing heat and vibration signals to predict abnormal conditions in wind turbines (Nabi et al., 2021).

3.5.2 Convolutional Neural Network (CNN)

Convolutional Neural Networks (CNN) are among the most effective algorithms for extracting complex patterns from data and are widely used in image and time signal processing, including wind turbine performance monitoring.

In this research, the CNN algorithm was used to analyze heat and vibration signals taken from the physical model in Simulink, after converting them into two-dimensional time arrays (2D feature maps) representing longitudinal and transverse changes in data over time(Chen et al., 2021).

The proposed CNN model consists of three main stages:

1- Convolution Layer: Filters of length k are applied to data vectors to extract local features, expressed mathematically as follows:

$$y_{i,j}^{(l)} = f\left(\sum_{m,n} x_{i+m,j+n}^{(l-1)} \cdot w_{m,n}^{(l)} + b^{(l)}\right) \quad (11)$$

- $y_{i,j}^{(l)}$, It is the output of the layer.
- $x_{i+m,j+n}^{(l-1)}$, it is the input of the previous layer.
- $w_{m,n}^{(l)}$, it is the filter weights.
- $b^{(l)}$, it is the bias coefficient (bias).
- $f(\cdot)$, activation function (usually ReLU).

2- Pooling Layer: Used to reduce spatial dimensions and minimize the number of operations while preserving the essential characteristics of the data, expressed as:

$$p_{i,j} = \max_{(m,n)\in R_{i,j}} y_{m,n} \quad (12)$$

Where $R_{i,j}$ is the specific aggregation area for each location (i, j).

3- Fully Connected Layer: Collects the features extracted from the previous layers and connects them to the final classification outputs. In this work, this layer is used to classify the condition into two categories: Normal or Faulty.

After training on 365 days of temperature and vibration data, the CNN model proved effective in capturing rapid changes in momentary vibrations and sudden temperature fluctuations, patterns that are often missed by traditional linear networks.

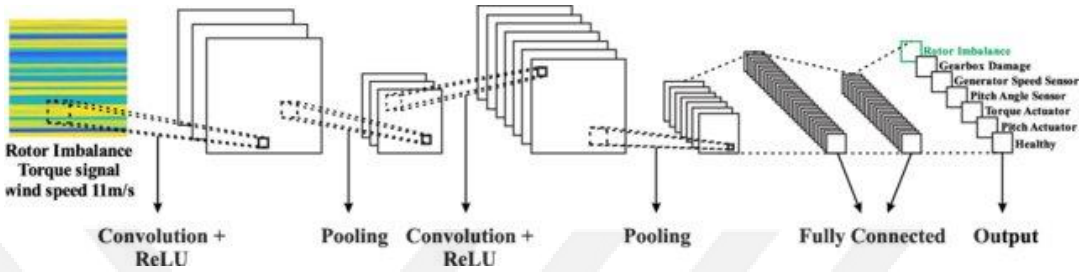


Figure 10 The general structure of the Convolutional Neural Network (CNN) used in classifying the operating conditions of wind turbines (Zare & Ayati, 2021).

3.5.3 Artificial Neural Network (ANN) Model

Artificial neural networks (ANNs) are among the simplest and oldest models of artificial intelligence used in complex signal analysis. These networks are modeled after the human brain, consisting of a group of interconnected neurons arranged in layers, including an input layer, hidden layers, and an output layer. Each neuron receives signals from the previous neurons and then processes them through an activation function to determine the appropriate outputs (Rudin et al., 2011).

In this research, an ANN model is used to classify signals measured from the turbine system such as temperature and vibration into specific operational categories (normal condition, overheating, excessive vibration, etc.). The goal of this model is to create a mathematical map that links sensor data to the actual state of the system without the need for complex physical models (An et al., 2011).

The neural network relies on a training process to adjust the internal weights between nodes in order to reduce the error between the expected and actual outputs (C. Zhang & Yang, 2023). This is done using the backpropagation algorithm, which updates the weights according to the following relationship:

$$w_{ij}(t+1) = w_{ij}(t) - \eta \frac{\partial E}{\partial w_{ij}} \quad (13)$$

- w_{ij}: The weight connecting cell i to cell j.
- η: Learning Rate.
- E: Error function, usually measured by Mean Squared Error, as in the equation:

$$E = \frac{1}{2}\sum\left(y_{output} - y_{target}\right)^2 \quad (14)$$

After training, the network becomes capable of predicting the future state of the system based on measured inputs.

In turbine applications, these models are used to predict early failures by analyzing subtle changes in measured values.

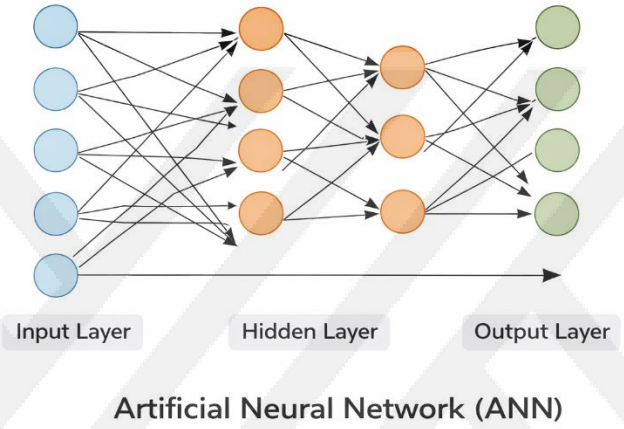


Figure 11 Simplified diagram of an ANN network for classifying the operating status of a wind turbine

3.5.4 Support Vector Machine (SVM) Algorithm

The SVM algorithm is one of the most effective methods for classifying high-dimensional data, especially in fault diagnosis applications in industrial systems such as wind turbines (Stetco et al., 2019).

The idea behind the algorithm is to find a hyperplane that separates different categories of data (such as normal and abnormal conditions) with the largest possible margin, thereby achieving high generalization capability for the model.

In this research, the SVM algorithm was used to classify operational data extracted from the turbine model in the MATLAB/Simulink environment, which includes temperature and vibration signals as the main inputs to the intelligent system.

The Radial Basis Function (RBF Kernel) was used to enable the model to handle nonlinear data resulting from sudden wind or mechanical load changes (B. Wang et al., 2025). The SVM model aims to find the optimal separator described by the following general equation:

$$0 = b + x \cdot w \tag{15}$$

Where: w eights Vector, x Feature Vector, b Bias

The goal of the algorithm is to maximize the margin between the two categories, which is formulated as an optimization problem:

$$min_{w,b}(1/2)||w||^2 \quad \text{subject to} \quad y_i(b + x \cdot w) \geq 1 \tag{16}$$

In the case of data that cannot be completely linearly separated, a kernel function is used to transform the data into a higher-dimensional space:

$$K(x_i, x_j) = \exp(-\frac{\|x_i - x_j\|^2}{2\sigma^2}) \tag{17}$$

Where σ is the width of the nuclear deal

$$K(x_i, x_j) = \exp\left(-\frac{||x_i - x_j||^2}{2\sigma^2}\right) \tag{18}$$

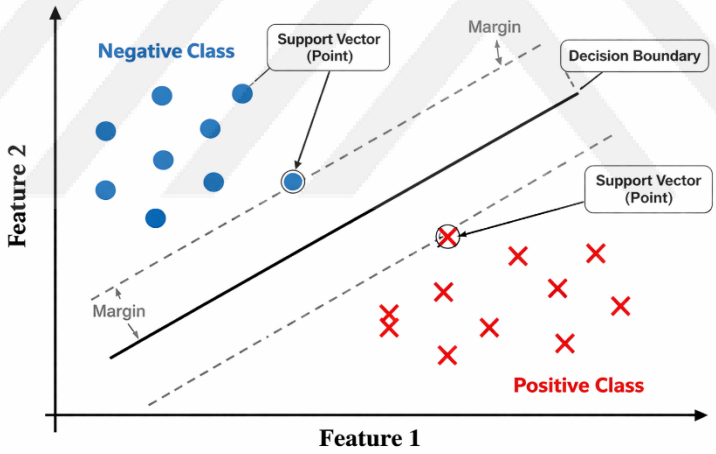


Figure 12 The conceptual structure of the SVM algorithm in classifying the operating conditions of wind turbines

3.5.5 Autoencoder Model

Autoencoder is a deep learning model designed to learn a compressed representation of data by reconstructing it. In this project, Autoencoder is used to identify abnormal conditions in wind turbines by comparing the reconstructed real signals with the original signals (Temperature & Vibration signals) (Malhotra et al.,

2016). Any significant deviation in the reconstruction is considered an indication of a system malfunction. The Autoencoder consists of three main stages:

- **Encoder:** Dimension reduction and extraction of features associated with the normal state.
- **Bottleneck (Latent Space):** A compressed representation that stores the essential characteristics of the signal.
- **Decoder:** Reconstruction of the original signal.

When a fault occurs, the model fails to accurately reconstruct the signal, resulting in a high reconstruction error value, which enables the system to detect faults even in the absence of obvious signs.

This is the formal Autoencoder equation:

$$h = f(W_e x + b_e)$$

$$\hat{x} = f(W_d h + b_d)$$

$$L(x, \hat{x}) = \sum (x - \hat{x})^2 \quad (19)$$

Where:

x: Original signal

$\hat{x}$: Reconstructed signal

h: Feature representation latent space

W_e, W_d, b_e, b_d: Model coefficients

f: Activation function (e.g., ReLU or sigmoid)

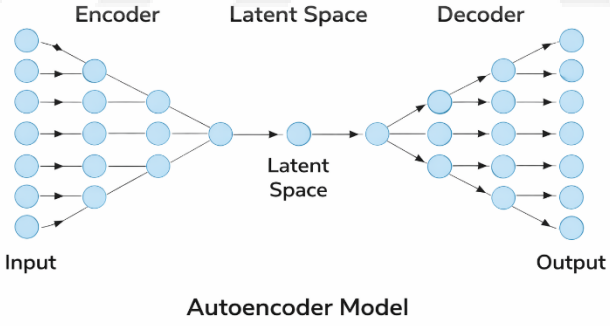


Figure 13 The proposed structure of the Autoencoder model used to detect anomalies in the temperature and vibration data of wind turbine

3.6 Implementation Process

This section describes the practical steps taken to implement the wind turbine anomaly prediction model, starting with the preparation of real wind data collected

over a full year in Istanbul, through the construction of the physical model in the Simulink environment, to the design of the deep learning system and the evaluation of its performance.

3.6.1 Data Acquisition

Actual wind data for the city of Istanbul for one full year (365 days) obtained from the Wolfram Mathematica 14.3 platform was used.

The main variables include: (Wind Speed (m/s), Temperature (°C), Vibration (m/s^2)).

The data was trimmed and converted to match the actual operating range of the turbine (3–25 m/s).

3.6.2 Building the physical model in Simulink

An integrated model of the turbine was created in MATLAB/Simulink, including: (A 12.3 kW wind turbine model, A permanent magnet synchronous motor (PMSM) generator, A DC-DC boost converter, MPPT unit for power control, Vibration and temperature sensors, Mechanical shaft inertia and damping, Thermal system and overcurrent Fault Injection).

PS-Simulink Converter + Ideal Sensors were added to read and inject temperature and vibration into the MATLAB environment.

3.6.3 Fault Injection

Various conditions were generated to simulate realistic scenarios:

(Condition Description, Normal operation, Fault-1 Gradual overheating, Fault-2 Sudden vibration pulse, Fault-3 Double overheating + vibration). Nonlinear degradation effects were added to show realistic deviation in the abnormal.

3.6.4 Data extraction and preparation Data Pre-processing

The following steps were applied: (Normalization min-max, Noise removal using Washout Filter, Data division into: 70% training, 15% testing, 15% validation).Conversion of signals into window segments for use in time networks (LSTM).

3.6.5 Training artificial intelligence models

The following models were trained:

(Target model, LSTM Time series anomaly detection, CNN Spectral pattern analysis, ANN Simple model for comparison, SVM Linear and nonlinear classification, Autoencoder Anomaly detection by reconstruction error). The Adam function was used for training, and MSE/Cross-Entropy for evaluation.

3.6.6 Integrating the intelligent system with Simulink

MATLAB models were linked as follows:

(Inputting live values from Simulink to MATLAB, calculating intelligent indicators (anomaly score), Sending an alert when a fault occurs).

The Real-Time Fault Alarm system was implemented using Triggered Subsystem in Simulink.

3.6.7 The final algorithm for the system

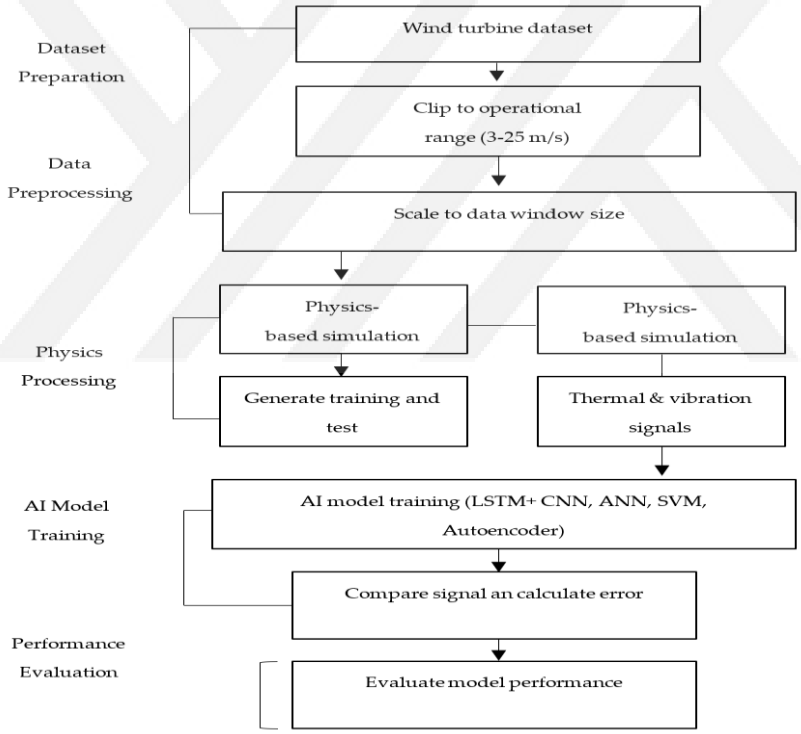


Figure 14 Diagram of the steps for implementing the proposed diagnostic system in the MATLAB/Simulink environment

3.7 Performance Evaluation

In this section, the performance of the proposed fault detection system will be evaluated based on a rigorous methodology covering statistical and technical aspects to ensure the reliability and practical applicability of the system in a wind turbine operating environment. The evaluation is based on analyzing the behavior of the two models used (SVM and LSTM) through real wind data, vibration signals, and temperature signals generated by the MATLAB/Simulink simulation environment.

The evaluation process includes the following indicators:

1- Accuracy is used to measure the ratio of correct predictions to the total number of samples. It is a general indicator of the model's effectiveness in distinguishing between normal and abnormal conditions.

2- Sensitivity (Recall) measures the system's ability to detect faults when they occur and is an important component in industrial systems that require early detection to avoid losses.

3- Specificity measures the system's ability to avoid false alarms during normal turbine operation, reducing unnecessary notifications and increasing the stability of the operating system.

4- Precision measures the ratio of correct alarms to the total number of alarms sent, and is an important metric in industrial safety systems to avoid unnecessary actions.

5- F1 score A composite measure that balances sensitivity and precision, providing a fair judgment of the model, especially in cases where the data is unbalanced.

6- Confusion Matrix Shows the details of the prediction between correct and false cases and is an important means of visually analyzing the results.

7- Model Performance Comparison The performance of the traditional model (SVM) is compared to the deep neural network model (LSTM) in terms of:

- Training speed
- Ability to learn from continuous time-series data
- Response to sudden spikes
- Accuracy in detecting abnormal signals, whether small or large

These metrics provide a clear picture of the system's efficiency and its ability to detect faults with high accuracy while minimizing false alarms. Detailed numerical results, tables, and graphs will be presented in Chapter 4.

3.8 Conclusion

This chapter reviewed the detailed steps for developing an intelligent system integrated with a wind turbine simulation environment for the purpose of early fault detection.

The work included multiple stages, starting with data collection and processing, through the construction of a realistic simulation model that integrates artificial intelligence with wind data, to the design of a prediction algorithm and the triggering of an alarm when an abnormal pattern is detected in the signal.

Advanced algorithms, including RBF kernel SVM and LSTM recurrent neural network models, were used to process time series and analyze signals. In addition, objective evaluation steps were included using statistical and technical metrics approved in industrial systems.

This chapter paves the way for the transition to Chapter 4, which deals with experimental results and their analysis, including graphs, signal behavior, model performance, and a comparison of the effectiveness of the developed systems, with a discussion of the model's ability to work in a real environment and potential challenges.

CHAPTER FOUR

SIMULATION AND SYSTEM IMPLEMENTATION

4.1 Introduction

This chapter aims to present the simulation environment used in the development and implementation of an early fault detection system for wind turbines, in addition to explaining the stages of data generation and preparation for training the artificial intelligence models used in this study. The proposed system is based on an integrated physical dynamic model of a 12.3 kW wind turbine within the MATLAB/Simulink environment, where normal operating conditions were simulated, and then gradual and sudden faults were introduced to evaluate the performance of the fault recognition mechanisms.

The model was designed to reflect the actual behavior of the turbine by integrating power generation mechanisms, blade angle control, and vibration and temperature analysis via virtual sensors. Realistic wind data taken from climate data for Istanbul for the year 2024 was also introduced to enhance the credibility of the experiments.

This chapter is a pivotal step in the research, as it represents the bridge between the theoretical foundation in the previous chapters and the experimental analysis and prediction models in the next chapter. Therefore, the chapter provides a detailed explanation of the model design process, the fault injection mechanism, and the methods of data collection and conversion into formats that can be used for training and evaluation.

4.2 Simulation Environment & Configuration

All simulations and data analysis were performed within the MATLAB R2025a /Simulink environment, due to its high capability to model complex dynamic systems and support integration with artificial intelligence algorithms. This software framework was chosen for its flexibility in dealing with renewable energy systems and the ability to connect the turbine model directly to artificial intelligence units within the MATLAB Workspace.

A 12.3 kW wind turbine was used in the simulation, consisting of:

- Wind turbine unit with performance coefficient Cp(λ,β)
- A mechanical system for the rotation axis and inertia mass
- A DFIG generator
- A power converter (Boost Converter)
- An MPPT controller to regulate the energy produced
- A Pitch Control system to adjust the angle of the blades

The following simulation settings were adopted:

- Solver: Backward Euler
- Step Time: 0.0001 s
- Simulation Mode: Discrete

This type of solver was chosen for its ability to handle nonlinear systems with rapid changes, while the short step time ensures accurate representation of instantaneous signals as vibration pulses.

Real wind data from the climate of Istanbul for 2024 was used, after trimming values outside the turbine's operating range (3–25 m/s).

Heat signals were generated from the Thermal Mass block, vibration signals from the Ideal Rotational Motion Sensor, and acceleration was derived using Discrete Derivative.

A direct link was established between the model and the MATLAB Workspace environment through the To Workspace and From Workspace blocks to export data for later analysis during AI model training.

4.3 Wind Turbine Model in Simulink

The wind turbine model was built within the Simulink environment in a realistic dynamic manner that reflects the physical behavior of wind energy systems. The model is based on the principle of converting the kinetic energy of wind into electrical energy through interconnected mechanical and electrical stages.

The system consists of the following main parts:

Wind Turbine Subsystem: Converts wind energy into mechanical torque based on the performance curve Cp (λ, β) derived from the relationship between wind speed, blade angle, and rotational speed.

Drive Train and Shaft Dynamics: Includes the rotation axis and mechanical inertia mass, where the interaction between the turbine and generator is accurately represented using differential rotation equations.

DFIG Generator Model: The double-fed induction generator is dynamically represented within Simscape Electrical, allowing for analysis of instantaneous power flow and system response to mechanical and electrical faults.

Power Conditioning Stage: Includes a Boost Converter to control and regulate the output voltage, as well as an MPPT (Maximum Power Point Tracking) algorithm to achieve the highest possible energy conversion efficiency.

Pitch Control System: Controls blade angles to maintain stability at high wind speeds, using a proportionally-integral (PI) controller that has been experimentally tuned to ensure smooth response.

This model is linked to the outputs of virtual sensors that continuously measure vibration and temperature, providing an accurate operational database that is later used to train machine learning algorithms.

4.4 Fault Injection

Fault injection was performed within the Simulink environment to study the response of the proposed system under abnormal operating conditions and analyze the effectiveness of artificial intelligence algorithms in early fault detection.

The faults were designed in a gradual and controlled manner to ensure realistic representation similar to field conditions in modern turbine systems.

The study adopted two types of basic faults:

Thermal Fault: A gradual increase in the temperature of the generator unit was applied to the Thermal Mass by 15–25% above the normal limit to simulate a cooling or overload fault. The resulting signal change was monitored and compared to reference values from normal operation.

Vibration Fault: A short-duration vibration pulse was injected into the signal of the Ideal Rotational Motion Sensor to simulate a fault in the rotational axis or dynamic balance. Noise was added within a limited range to test the models' ability to distinguish between natural disturbance and actual failure.

The simulation scenario included four main operating conditions: (Normal operation, thermal fault condition, vibration fault condition and combined fault condition)

Data for each case was collected chronologically over a complete operating cycle to enable balanced training of the AI models.

Figure 15 shows a graphical example of the change in temperature and vibration when a fault occurs. These signals were subsequently used in the data processing and analysis phase.

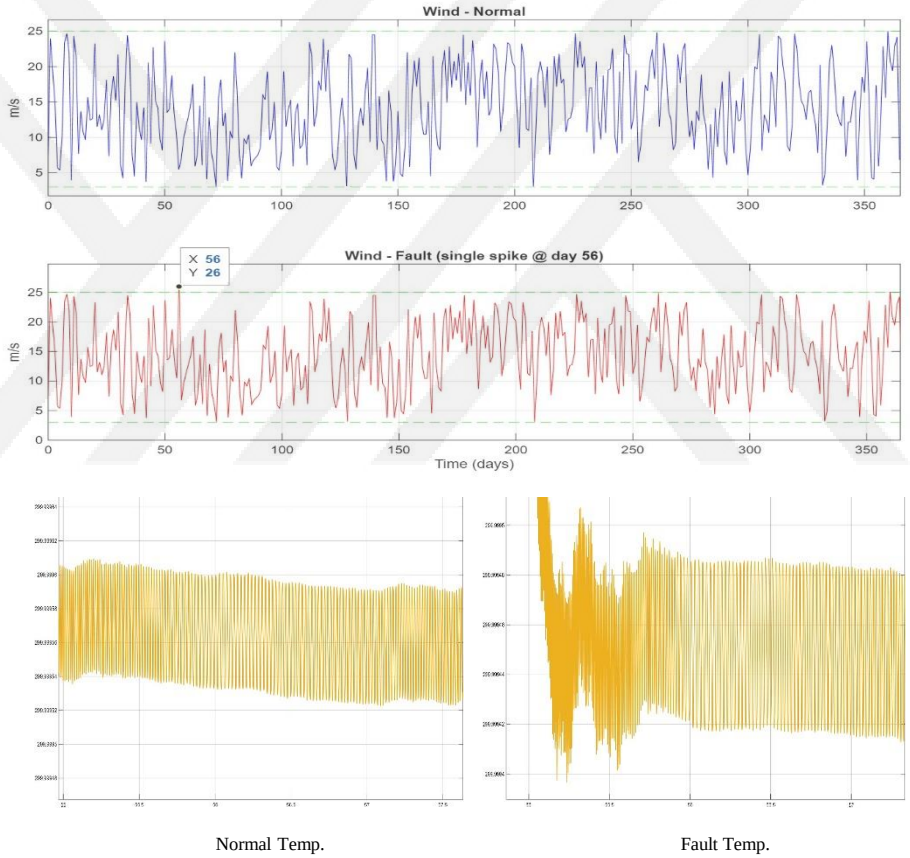


Figure 15 Response of thermal and vibration signals during fault injection in a physical model of a wind turbine.

4.5 Data Acquisition

The MATLAB/Simulink environment was used as the primary platform for capturing the turbine's operational signals during normal and fault conditions.

The system relied on virtual sensors within the physical model to produce accurate data representing the actual dynamic behavior of the turbine. The data acquisition process included the following steps:

recording vibration signals using the Ideal Rotational Motion Sensor element was used to measure angular velocity and displacement. Acceleration was calculated using Discrete Derivative to obtain a signal sensitive to sudden changes during faults. The signal showed clear sharp peaks at the moment of the fault (Figure 15 – Vibration section). Temperature recording the temperature was simulated using a Thermal Mass block connected to the generator and backplate. The signal showed thermal stability during normal operation, while thermal distortions appeared during the fault, as shown in Figure 15. Data storage in MATLAB Workspace. The following tools were used:

(To Workspace, From Workspace, MATLAB Scripts) to record signals in the following formats: (Array, Timeseries, Struct) and at a fixed sampling rate of 0.0001 sec to ensure high-frequency signal accuracy. Preparing training files. After collecting the data, the following was performed dividing the data into sliding windows suitable for LSTM networks. Saving the signals classified as Normal/Fault. Creating a dataset ready for training SVM, Autoencoder, and LSTM models.

4.6 Data Preparation for Artificial Intelligence

After collecting the turbine's operational signals, a series of steps were carried out to process the data and convert it into a format suitable for training artificial intelligence models. This stage aimed to remove noise, normalize the signals, and divide them into time windows that the models could learn from accurately. Signal Cleaning digital processing tools were used to remove interference caused by: (Momentary wind disturbances, Sample rate changes, Virtual sensor noise)

This included: Washout Filter to remove drift from vibration signals. Moving Average Filter to smooth temperatures without affecting the true thermal response. Data normalization to ensure training stability, the following were applied: (Min–Max Scaling to vibration signals, Z-Score Normalization to temperature data, The goal is to

bring all signals within a fixed numerical range that allows models to learn without bias.

Labeling the data was classified according to two main cases: (Normal operation, Fault Thermal or vibration fault). The fault classification was based on: (Exceeding the predefined thermal threshold, The appearance of an abnormal vibration peak on days 10/56/208 As shown by the simulation results, High reconstruction error for the Autoencoder model)

Sliding Window Segmentation since LSTM and Autoencoder deal with time series data, the signals were divided using fixed-length windows: (Window length: 64 samples, Overlap ratio: 50%). This preserves the continuity of temporal information. Dividing the data into training and testing sets the data was divided as follows: (70% for training, 15% for validation, 15% for testing).

Ensuring that each set contained both normal and abnormal cases (Stratified Split). Creating the final training files the data was saved in MATLAB files in the following formats: (. mat, Struct Array, Timeseries) to be ready for use in: (SVM, LSTM Autoencoder, Threshold-based Detector).

4.7 Summary

This chapter presents the practical basis on which this research methodology is based, as it builds an integrated dynamic simulation model of a 12.3 kW wind turbine within the MATLAB/ Simulink environment, including aerodynamic, mechanical, and electrical modeling, as well as the integration of virtual sensors to monitor temperature and vibration under normal and fault conditions.

The chapter demonstrated a clear scientific methodology for generating operational data using real wind data for the year 2024 and calibrating it to match the turbine's operating range. Multiple failure scenarios were designed, the most important of which were: (Thermal drift, Vibration spikes, which were injected realistically into the model to test the ability of AI systems to detect them early.

The chapter also addressed the techniques used to clean, normalize, and divide signals into time windows ready for deep learning, which is an essential step to ensure the quality of inputs before training models. An integrated database was created that

included both Natural / Faulty cases, allowing for an accurate comparison in Chapter 5 between the actual performance of each algorithm.

This chapter is the turning point of the thesis, as it combines: (Documented numerical modeling, Realistic data, Controlled fault injection, Processing of organized and balanced data for machine learning).

This lays the foundation for the analysis of results in Chapter 5, which will demonstrate the effectiveness of Autoencoder, SVM, ANN, and CNN-LSTM models in detecting and separating thermal and vibration faults before they turn into major turbine failures.

CHAPTER FIVE

RESULTS AND AI-BASED FAULT DIAGNOSIS

5.1 Introduction

This chapter presents the detailed results of the proposed system for early detection of faults in wind turbines through analysis of signals generated by a dynamic simulation model built in the MATLAB/Simulink environment. This chapter is a direct practical extension of the methodology discussed in the previous chapter, moving from the design and configuration phase to the testing and evaluation phase based on real signals generated from the physical model of the turbine. This phase aims to understand the operational behavior of the system under normal and abnormal operating conditions and to detect thermal and vibration changes that indicate an actual or imminent malfunction.

The chapter begins by presenting the operational response of the turbine during normal operation, showing the stability of the signals and the absence of abnormal patterns. Next, fault conditions resulting from temperature spikes or sudden vibration peaks within specific time intervals are analyzed, accurately reflecting the impact of mechanical and thermal failures on turbine components. The results of signal processing are then reviewed and converted into data suitable for use in artificial intelligence models.

The chapter also evaluates the performance of a set of algorithms trained on the collected data, including Autoencoder, SVM, ANN, and CNN-LSTM. The results of these models are analyzed in terms of predictive power, classification accuracy, and model response to anomalies, in addition to a comprehensive comparison of the performance of different models to determine the most effective method for detecting faults before they worsen and affect turbine efficiency.

5.2 Simulation Results

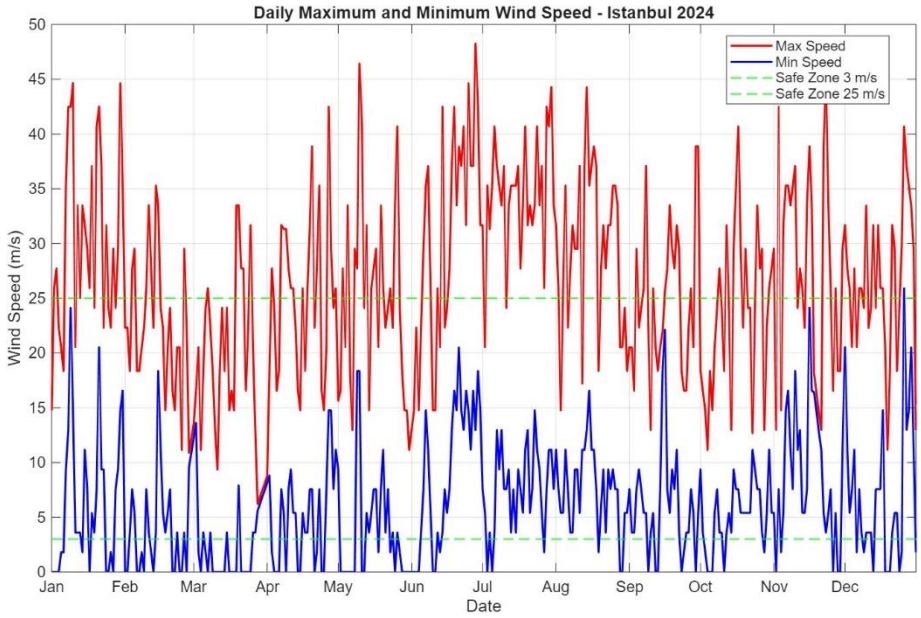


Figure 16 Raw wind signal before clipping (Istanbul – 2024)

This figure 16 represents the daily wind speed signal for a full year (2024) as imported from the climate database. Significant natural variability in speed can be observed within the range (0–50 m/s), with high peaks that cannot be directly used in the turbine model. Therefore, this signal was adopted as a preliminary input before being processed and filtered to obtain the safe operating range.

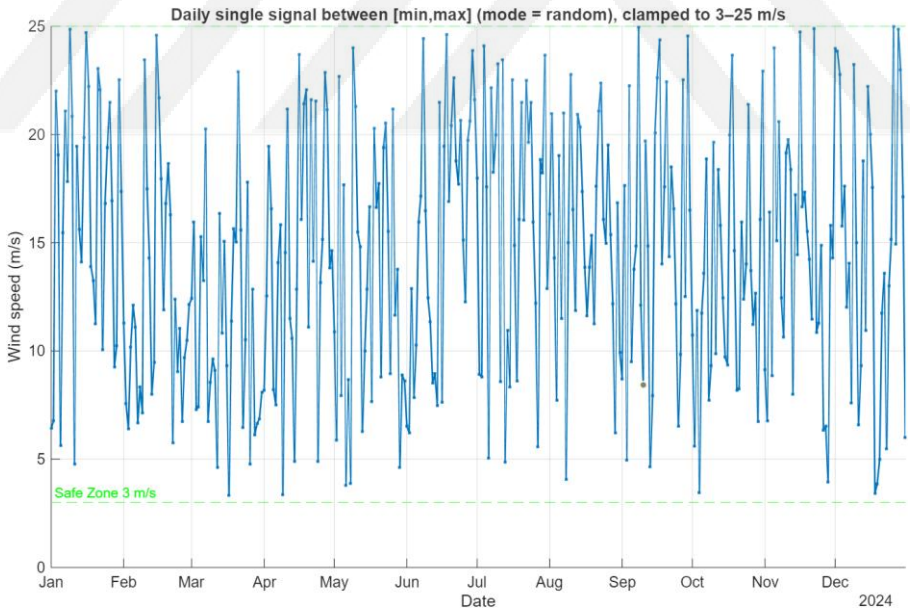


Figure 17 Wind signal after clipping to the operating range (Safe Operating Zone: 3–25 m/s)

After clipping the original signal to the turbine operating range (3–25 m/s), the signal became more representative of the industrial reality. The figure shows the removal of values below the critical speed of 3 m/s or above 25 m/s, to ensure that the physical model operates within safe limits and does not enter areas of shutdown or structural exposure.

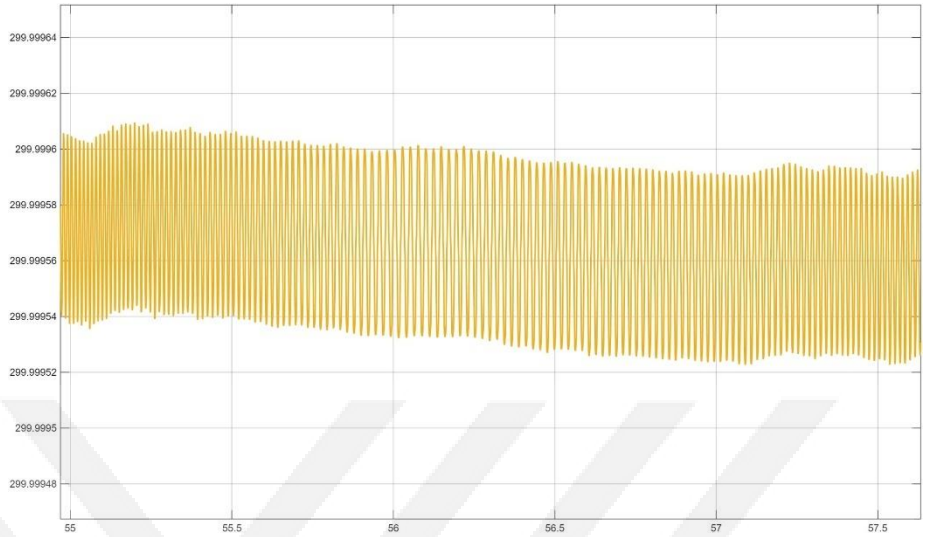


Figure 18 Temperature signal during normal operation (Normal Temperature)

The figure shows the temperature response in a normal case without any malfunction. The signal is stable and smooth, with consistency in frequency and amplitude, reflecting thermal equilibrium in the generator and mechanical axis. This behavior is considered the baseline for deep learning models that recognize the (normal context) before detecting deviations.

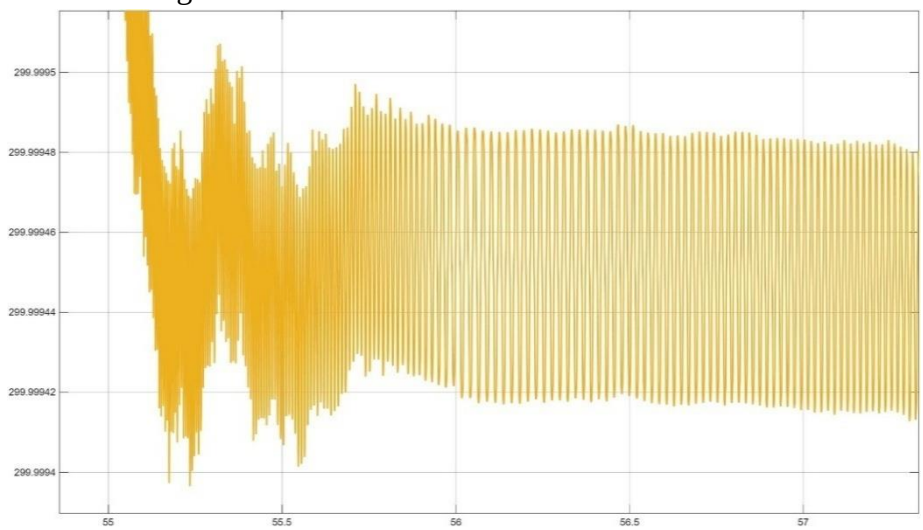


Figure 19 Temperature signal during a fault condition (Fault Temperature – Drift/Instability)

This figure shows an abnormal thermal state resulting from a gradual thermal fault. Distortions and a noticeable increase in amplitude compared to the normal state can be observed. This deviation is the factor on which the Autoencoder algorithm relies to measure the reconstruction error and distinguish the abnormal state.

5.2.1 Normal Operation

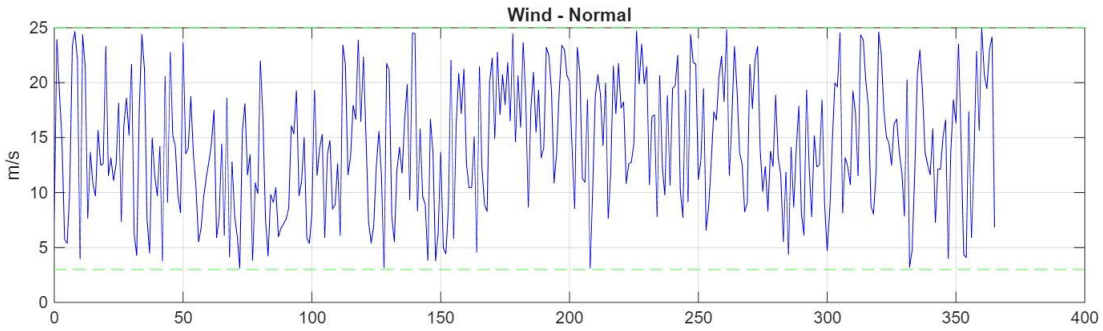


Figure 20 Normal Vibration Signal

Figure 20 shows the time for vibration signal over a full year in a normal state before any fault occurred. The signal range is between 3–25 m/s, which is the safe operating range for the turbine, as indicated by the safe zone limits with the green dotted line at the top and bottom of the figure.

The signal shows fluctuating behavior due to daily seasonal changes, but all values remain within acceptable operating limits without recording abnormal peaks or sudden interruptions. This behavior is considered a baseline that is later used to compare malfunctions and to evaluate the response of artificial intelligence algorithms in detecting deviations.

This figure also represents the input source for the turbine model in Simulink, where each time point was used as a wind speed value to calculate torque, power, and secondary signals such as heat and vibration. The absence of distortions or sharp peaks in this Graphic confirms that the model is in a stable normal operating condition.

5.2.2 Fault Operation

This section reviews the dynamic behavior of the turbine system during a fault and shows the effect of sudden increases or gradual deviations in operating variables such as wind speed, temperature, and vibration. This behavior represents the “simulated operating environment” environment in which early fault detection algorithms operate, and is the basis on which performance accuracy is evaluated.

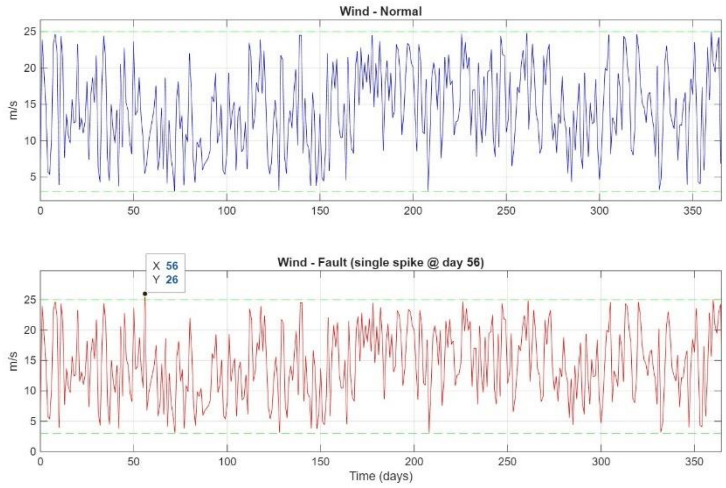


Figure 21 Wind signal during fault (Fault @ Day 56)

The figure shows an abnormal spike on day 56, where the speed rose above the upper limit of the safe operating range (25 m/s). This momentary increase represents a dynamic fault in the operating conditions, leading to a sudden disturbance in the generated mechanical power and aerodynamic torque, which is subsequently reflected in the rest of the thermal and vibration signals.

This situation is considered a transient fault, which is the most challenging for early detection systems because it is extremely short and may not be detected by conventional techniques based on time averages.

5.3 Feature Analysis

5.3.1 Introduction to Extracted Features

In order to characterize turbine behavior and distinguish between normal and abnormal conditions, a set of statistical indicators was extracted from temperature and vibration signals. These features were chosen because they are computationally simple and widely used in the diagnosis of faults in rotating systems. They include the following:

Root Mean Square (RMS): Reflects the overall level of signal energy and is an indicator of thermal and mechanical power or load.

Standard deviation (STD): Measures the degree of signal dispersion around the mean. A high STD value typically indicates increased fluctuation and instability.

Maximum value (Max): Represents the highest value recorded within the signal and is sensitive to sharp peaks or sudden disturbances.

Peak-to-Peak: The difference between the highest and lowest values, reflecting the severity of signal fluctuation.

Energy: The sum of the squares of the values, representing the total energy transferred during the measurement window.

These indicators were calculated for both heat and vibration signals, in both normal and fault conditions, for three representative days: day 10, day 56, and day 208. Figure 22 (a-b-c) show a visual summary of the differences between the two conditions.

Although CNN models are typically designed for raw signal processing, in this study they are applied to structured feature representations to capture local patterns and interactions between extracted statistical features. This allows the CNN layer to learn higher-level abstractions from correlated feature sets, improving the model's ability to distinguish between normal and faulty conditions. The combination with LSTM further enables capturing temporal dependencies across consecutive feature windows.

The hyperparameters of the models were selected empirically based on preliminary experiments to achieve stable training performance. For the CNN-LSTM model, key parameters included the number of convolutional filters, kernel size, LSTM units, batch size, learning rate, and number of training epochs. These parameters were kept consistent across experiments to ensure fair comparison between models.

To improve methodological transparency and reproducibility, the main hyperparameters used for training the proposed AI models are summarized in Table 2.

Table 2 Hyperparameters of the Proposed AI Models.

Model	Hyperparameter	Value
Autoencoder	Hidden latent dimension	32
Autoencoder	Activation function	ReLU
Autoencoder	Loss function	Mean Squared Error
Autoencoder	Optimizer	Adam
SVM	Kernel type	RBF
SVM	Kernel scale σ	Auto
SVM	Box constraint (C)	1
ANN	Hidden layers	2
ANN	Hidden neurons	64, 32
ANN	Activation function	ReLU
ANN	Output activation	Softmax
ANN	Optimizer	Adam
CNN-LSTM	Filters	32
CNN-LSTM	Kernel size	3
CNN-LSTM	LSTM units	64
CNN-LSTM	Batch size	128
CNN-LSTM	Learning rate	0.001
CNN-LSTM	Epochs	200

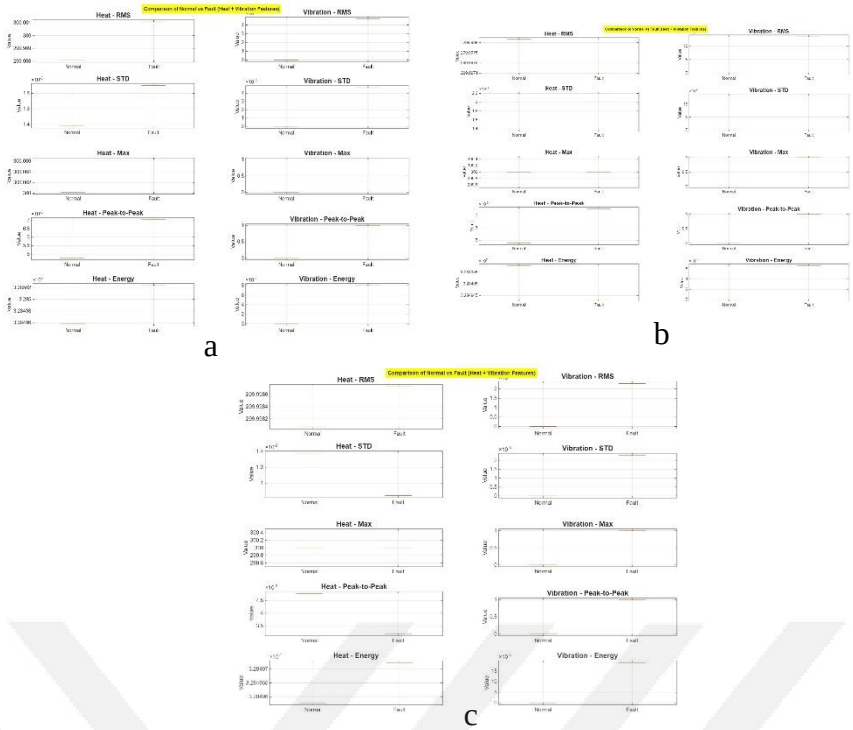


Figure 22 Comparison of heat and vibration characteristics for days (10,56,208)

Table 3 Statistical characteristics of the (Heat Signal)

Day	RMS_ N	STD_ N	Max_ N	P2P_ N	Energy_ N	RMS_ F	STD_ F	Max_ F	P2P_ F	Energy_ F
10	300	0.001379	300	0.0047779	3.285×10^{7}	300	0.0019011	300.01	0.0070324	3.285×10^{7}
56	300	0.001379	300	0.0047779	3.285×10^{7}	300	0.0021911	300	0.0075302	3.2849×10^{7}
208	300	0.001379	300	0.0047779	3.285×10^{7}	300	0.00084314	300	0.0031805	3.285×10^{7}

The table 3 shows a detailed comparison between the statistical values for the normal state and the fault state on days 10, 56, and 208. It can be seen that the RMS and Energy values remain almost constant on all days, indicating that the fault did not affect the total power or energy level of the thermal signal.

However, a clear increase in standard deviation (STD) and peak-to-peak (P2P) values is observed during the fault condition, especially on days 10 and 56, reflecting increased dispersion and sudden changes within the signal. These increases are an important indicator of an abnormal pattern despite the stability of the energy,

confirming that the anomaly was behavioral rather than a defect in the thermal power level.

On day 208, a clear decrease in the STD value was observed compared to the rest of the days, indicating that the anomaly here is less severe or gradual in nature rather than short shocks.

Table 4 Statistical Properties of (Vibration Signal)

Day	RMS_ N	STD_ N	Max_ N	P2P_ N	Energy_ N	RMS_ F	STD_ F	Max_ F	P2P_ F	Energy_ F
10	0	0	0	0	0	0.00046816	0.00046816	1	1	8×10^{-5}
56	0	0	0	0	0	0.0013447	0.0013447	1	1	0.00066
208	0	0	0	0	0	0.0022816	0.0022815	1	1	0.0019

The table 4 shows that the normal signal on almost all days is close to zero in terms of statistical values, which is logical because the vibration in normal operating conditions is very low.

In the case of a malfunction, however, there is a gradual and clear increase in RMS, STD, and Energy over days 10, then 56, then 208. This behavior indicates that the severity of the mechanical fault increased over time, and that the vibrations changed from a minor fault on day 10 to a more pronounced and severe pattern on day 208.

The values of P2P and Max (approximately equal to 1 in the case of a defect) confirm that the signal contains sharp spikes, which are one of the clearest signs of a mechanical problem in the turbines or the rotation system.

Thus, the vibration signal shows greater sensitivity in detecting faults than the temperature signal, confirming its importance as a key diagnostic variable.

In the case of malfunction, a gradual and clear increase in RMS, STD, and Energy appears over days 10, then 56, then 208. This behavior indicates that the severity of the mechanical fault increased over time, and that the vibrations changed from a minor fault on day 10 to a more pronounced and severe pattern on day 208.

The values of P2P and Max (approximately equal to 1 in the case of a defect) confirm that the signal contains sharp spikes, which are one of the clearest signs of a mechanical problem in the turbines or the rotation system.

Thus, the vibration signal shows greater sensitivity in detecting faults than the temperature signal, confirming its importance as a key diagnostic variable.

5.3.2 Analysis of characteristics on day 10

Figure (22-a) shows a comparison of the statistical characteristics of the normal and fault conditions on day 10.

For the heat signal, the RMS and Energy values appear to be approximately constant between the two cases (RMS ≈ 300, Energy ≈ 3.285×10^{7}), indicating that the overall average heat is not significantly affected by the onset of the defect.

On the other hand, the characteristics associated with dispersion increase significantly. The standard deviation increases from about 0.001379 in the normal state to 0.001901 in the disturbance state, and the peak-to-peak value increases from 0.00478 to about 0.00703. This indicates that the imbalance in its early stages translates into higher thermal vibrations around the same average level.

As for the vibration signal, the normal condition shows almost zero values, while the fault condition shows small but clear vibrations (RMS ≈ 4.6×10^{-4}, Max = 1, P2P = 1, energy ≈ 8×10^{-5}).

This represents an expected pattern for the onset of mechanical malfunction, where very weak vibrations appear that are not present in the normal state.

5.3.3 Analysis of characteristics on day 56

Day 56 represents a more advanced state of failure. As shown in Figure (22-b), the RMS and Energy values for heat remain approximately constant, but the STD and Peak-to-Peak values increase significantly compared to day 10, indicating more pronounced thermal disturbance.

The vibration signal shows clear differences between the two cases, with RMS and STD values rising to about 1.34×10^{-3} and clear single peaks (Max = 1, P2P = 1) appearing, with higher energy (about 6.6×10^{-4}).

This increase reflects that the imbalance has become stronger at the mechanical level, so that its effect is clearly visible on the vibrations, in parallel with greater thermal fluctuation.

5.3.4 Analysis of characteristics on day 208

On day 208, Figure (22-c) shows a pattern that is completely different from the previous two days.

In the temperature signals, the RMS and Energy values remain almost constant, while STD and Peak-to-Peak decrease slightly in the fault case compared to the normal case. This indicates that the fault no longer affects the thermal behavior and that the temperature has become more stable during the fault period.

However, vibration shows a very clear increase compared to the normal state: RMS $\approx 2.28\times10^{-3}$, STD $\approx 2.28\times10^{-3}$, with clear peaks (Max = 1, P2P = 1), and energy $\approx 1.9\times10^{-3}$.

This represents a case of purely mechanical imbalance, which is fully manifested in the vibrations but not in the heat.

5.3.5 General discussion and engineering implications

The most important conclusions from the analysis of the characteristics can be summarized as follows:

RMS and Energy are insensitive to the fault They remain approximately constant over the three days, which means that the fault does not change the overall power or heat load, but only changes the signal pattern.

STD and Peak-to-Peak are the most discriminating indicators. They rise clearly on days 10 and 56, reflecting increased thermal disturbance.

Vibration is the primary carrier of mechanical instability. On days 56 and 208, the instability appears strongly in the vibrations and becomes more pronounced as the condition progresses.

These characteristics will form the basis for building artificial intelligence models later, as Autoencoder relies on differences in STD and P2P, and SVM, ANN, and CNN-LSTM rely on the complete set of features.

5.4 Results of artificial intelligence models for fault detection

This section presents the results of applying different artificial intelligence models for turbine fault detection, namely: Autoencoder, Support Vector Machine, Artificial Neural Network, and the hybrid CNN-LSTM model. The performance of each model was evaluated using a confusion matrix and ROC curve, in addition to quantitative performance indicators: Accuracy, Precision, Recall, F1-score, and AUC.

Figure (23) illustrates the convergence behavior of the proposed CNN–LSTM model during the training process, where both the RMSE and loss values decreased gradually and steadily across the number of epochs, indicating the stability of the learning process and the effectiveness of the model optimization and RMSE and loss convergence were used as part of the process to verify model stability during training.

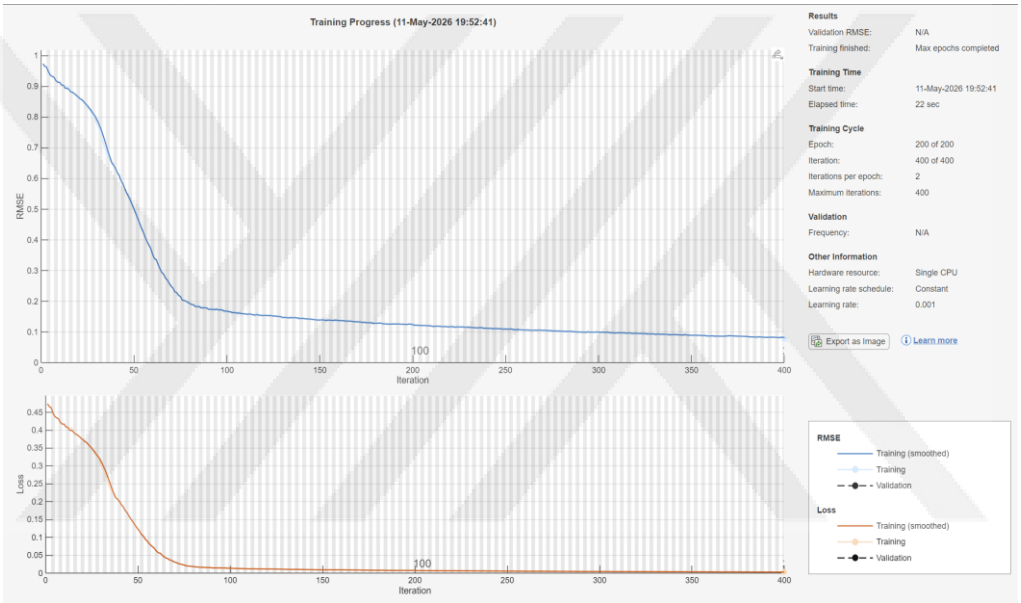


Figure 23 The convergence behavior of the CNN–LSTM model is illustrated by the decrease in RMSE and loss values during training.

The reconstruction error was calculated by determining the absolute difference between the original signal and the signal reconstructed by the proposed CNN–LSTM model, according to the following equation:

$$Error = |x - \hat{x}| \tag{20}$$

where (x) represents the original signal input to the model, and $(\hat{x})$ represents the reconstructed signal output from the model. Subsequently, the reconstruction error values for all samples were statistically analyzed to determine the detection threshold.

The results showed that the lowest reconstruction error value was 0.0202, the highest was 0.1514, and the average error was 0.0772. Based on the error distribution under normal operating conditions, a threshold value of 0.1416 was empirically selected as the upper acceptable limit for normal conditions.

Samples that exceeded this threshold were classified as abnormal or potential faults. The results showed that only 3 out of 303 samples exceeded the specified threshold value, indicating the model's stable performance and a low false alarm rate under various operating conditions as shown in Table 5.

Table 5 Quantitative analyses of the reconstruction error sensitivity of the CNN–LSTM model.

Parameter	Value
Minimum reconstruction error	0.0202
Mean reconstruction error	0.0772
Maximum reconstruction error	0.1514
Selected detection threshold	0.1416
Samples exceeding threshold	3 / 303

5.4.1 Autoencoder-Based Anomaly Detection

The Autoencoder model was used as an unsupervised learning method to detect anomalies in heat and vibration signals. The model is based on the principle of signal reconstruction and identifies anomalies based on the reconstruction error compared to a specified threshold. The model was trained using only normal operating data and then tested on data containing both normal and abnormal cases.

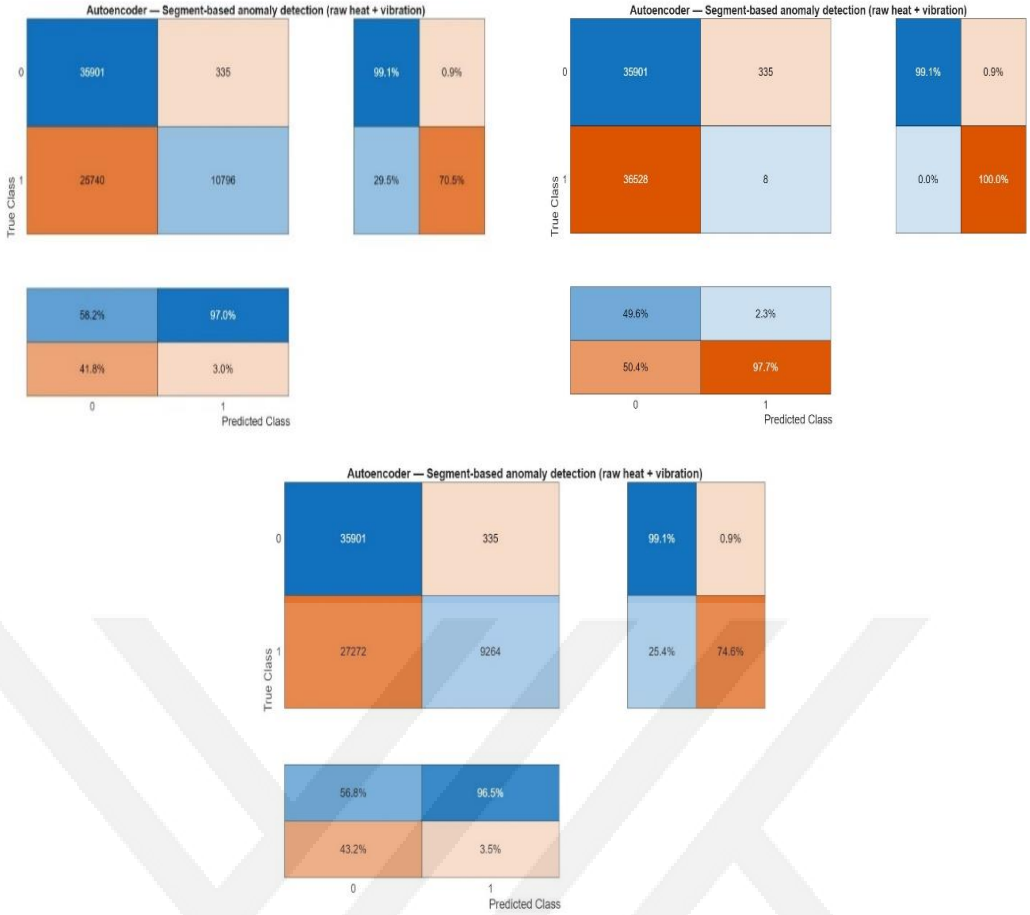


Figure 24 shows the confusion matrix results for the Autoencoder model for days (10, 56, 208)

It can be seen that the model performs well in classifying normal cases, with very high True Negatives values on all days. In contrast, the model fails to detect a large number of abnormal cases, as evidenced by the clear increase in False Negatives values, especially on days 56 and 208.

On day 10, the model was able to detect a limited number of anomalies, but a large proportion of them were classified as normal cases. This behavior worsens on day 56 and reaches its worst on day 208, where the model is almost completely unable to detect anomalies.

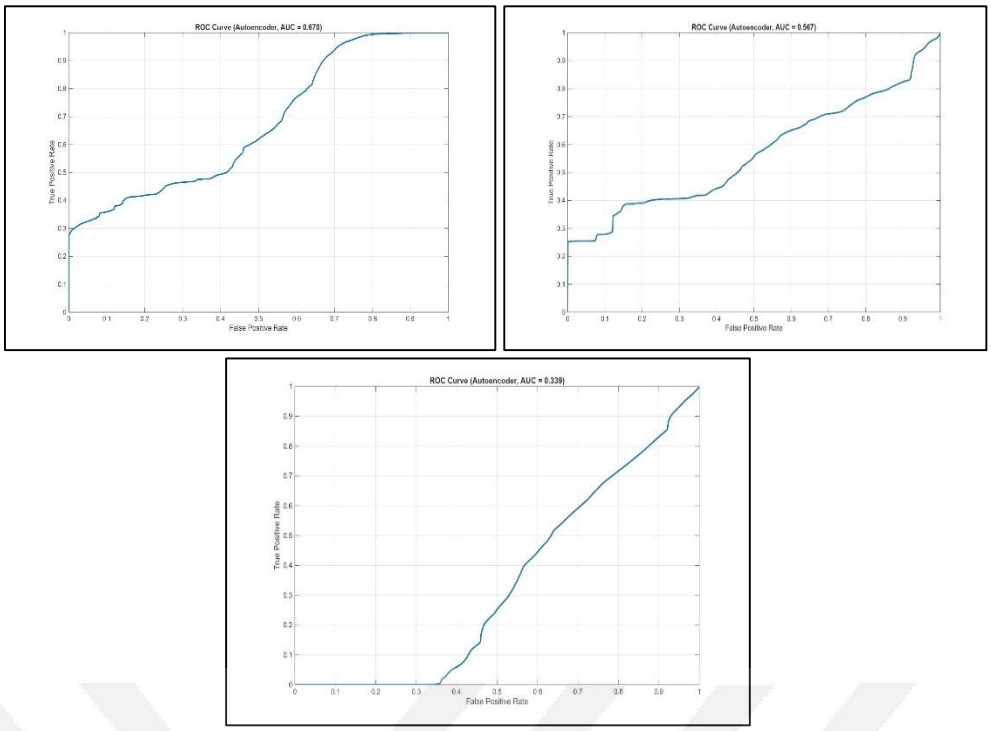


Figure 25 shows the ROC curves for the Autoencoder model for the three days

On day 10, the AUC value reached 0.670, which indicates a weak to moderate separation ability between the two cases. On day 56, the AUC value dropped to 0.567, indicating a clear deterioration in the model's performance.

On day 208, despite the apparent high AUC value, this value does not reflect good actual performance, but rather the complete collapse of the model's ability to classify anomalies, which is confirmed by other performance indicators. The low values for precision, recall, and F1 score indicate that the selected decision threshold failed to effectively distinguish between normal and abnormal samples, resulting in poor practical performance of the classification process.

Table 6 Performance metrics of Autoencoder anomaly detection model

Day	Accuracy	Precision	Recall	F1-score	AUC
10	0.6417	0.9699	0.2955	0.4530	0.6704
56	0.6206	0.9651	0.2536	0.4016	0.5674
208	0.4934	0.0233	0.0002	0.0004	0.3390

The results show that the Autoencoder model suffers from a clear weakness in detecting anomalies, despite achieving high Precision values on some days. This is because the model relies solely on normal operating data during training, making it unable to effectively distinguish between normal and abnormal patterns when the differences are subtle or gradual.

The sharp decline in Recall and F1-score values, especially on day 208, indicates that the model fails to detect actual malfunctions, making it unsuitable for use as a standalone diagnostic system in this application. Furthermore, the instability of performance across different days limits its reliability in real-world operating environments.

Based on this, the Autoencoder model can be considered a useful preliminary exploratory tool for monitoring general deviations in the data, but it is not sufficient to rely on it as the primary model for fault detection, which calls for a transition to more sophisticated monitoring models, as will be explained in the following sections.

5.4.2 SVM-Based Anomaly Classification

A radial basis function (RBF) Support Vector Machine (SVM) algorithm was used as a supervised learning method to classify operational states into normal and abnormal states, based on characteristics extracted from heat and vibration signals. This model aims to improve detection capability compared to the Autoencoder model, especially in cases of gradual and complex anomalies.

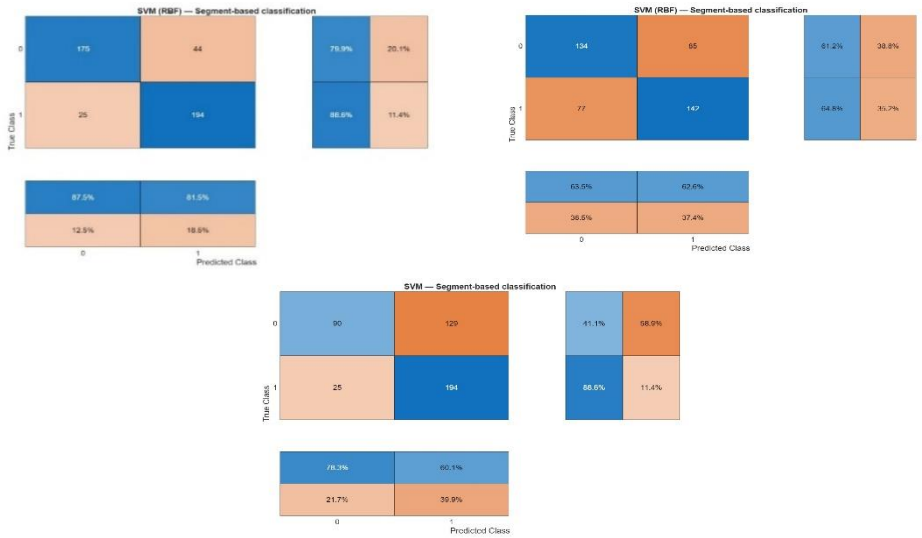


Figure 26 shows the confusion matrices for the SVM model for days (10, 56, 208)

The confusion matrix results show a clear improvement in the SVM model's ability to detect malfunctions compared to the Autoencoder model. On day 10, the model achieved a good balance between detecting normal and abnormal cases, with a clear decrease in False Negatives, indicating the model's effectiveness in capturing actual failures.

On day 56, a relative decline in performance was observed, but the model maintained acceptable classification ability compared to the unmonitored model. On day 208, the SVM model showed better stability in detecting anomalies compared to the Autoencoder, despite significant operational changes in the data.

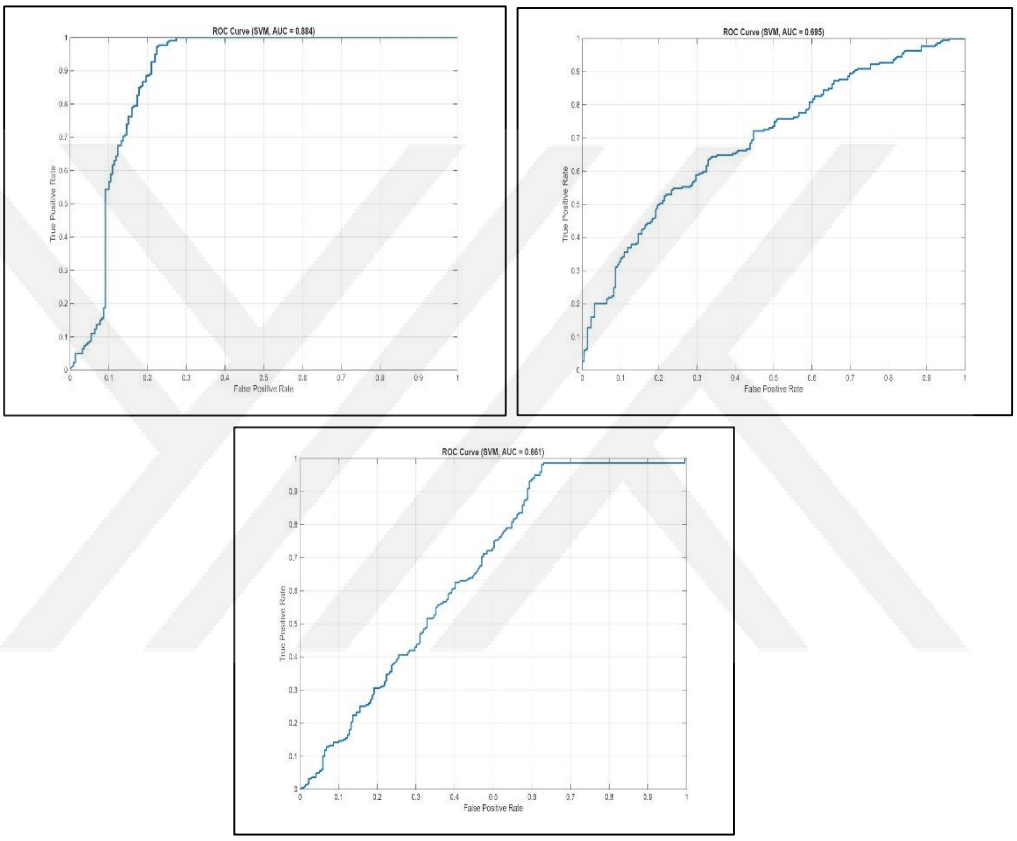


Figure 27 shows the ROC curves for the SVM model for the three days

The ROC curves showed a significant improvement in the ability to distinguish between normal and abnormal cases using the SVM model, with an AUC value of approximately 0.884 on day 10, indicating high discrimination ability. Although the AUC value decreased on days 56 and 208, it remained within an acceptable range, reflecting relative stability in model performance over time compared to unmonitored models.

Table 7 Performance metrics of SVM-based anomaly detection model

Day	Accuracy	Precision	Recall	F1-score	AUC
10	0.8425	0.8151	0.8858	0.8490	0.8840
56	0.6301	0.6256	0.6484	0.6368	0.6954
208	0.6484	0.6006	0.8858	0.7159	0.6612

The quantitative results indicate that the SVM model clearly outperforms the Autoencoder model in detecting anomalies, especially in terms of Recall and F1-score values, reflecting a higher ability to capture true anomalies and reduce false negatives. This improvement is attributed to the nature of supervised learning, where the model benefits from anomaly data during training, unlike unsupervised models.

Based on this, the SVM model can be considered a more reliable candidate for use in early diagnosis systems, noting that its performance may be affected by changes in data distribution over time, which calls for a transition to deeper and more adaptive models, as will be reviewed in subsequent sections.

5.4.3 Artificial Neural Network (ANN) Results

An Artificial Neural Network (ANN) model of the Multi-Layer Perceptron (MLP) type was applied as a supervised learning algorithm to classify operating conditions into normal and fault conditions, based on statistical characteristics extracted from temperature and vibration signals. The use of ANN aims to improve fault detection accuracy compared to the Autoencoder model, especially in cases that require direct discrimination between the two categories.

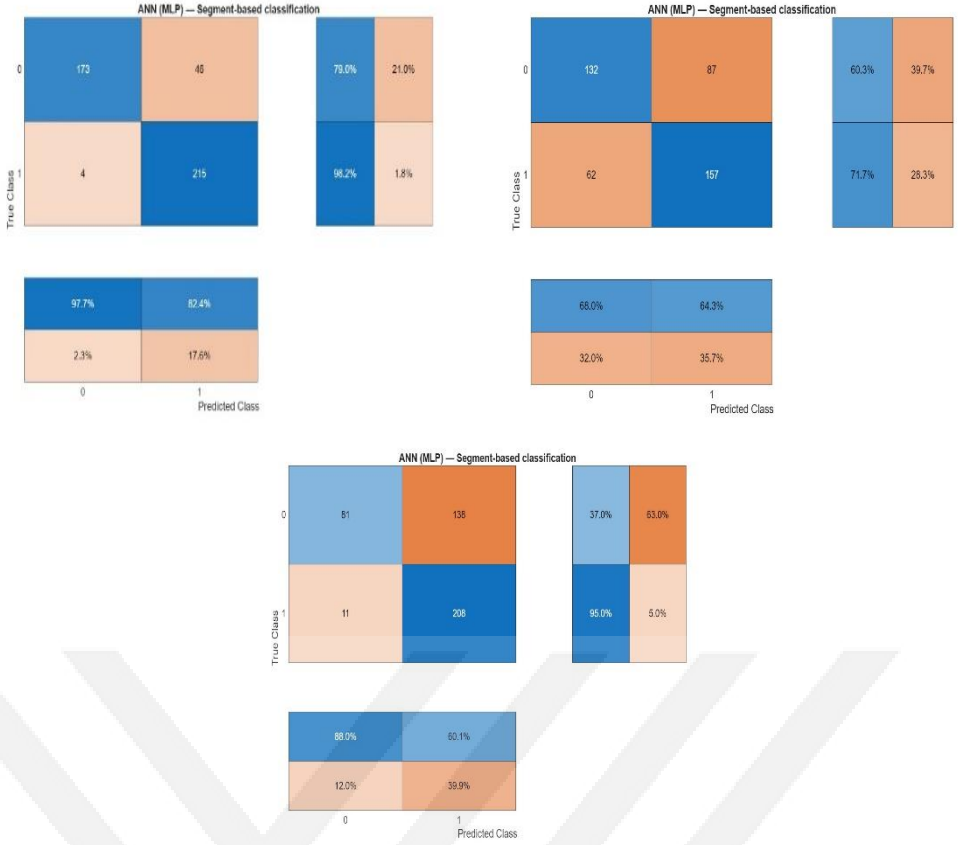


Figure 28 shows the confusion matrices for the ANN model for days (10, 56, 208)

Figure 28 shows the confusion matrices for the ANN model for days 10, 56, and 208. On day 10, the model showed high performance in detecting faults, with a large number of correct values recorded for both normal and fault conditions and a significant decrease in misclassification, reflecting the model's strong ability to separate the two classes. On day 56, the model's performance declined relatively, with an increase in confusion between the two categories, indicating that the model was affected by a change in the signal pattern or an increase in the similarity between normal and abnormal operating characteristics. On day 208, the model showed a high ability to detect malfunctions, but this was accompanied by an increase in false alarms, indicating that the model tended to prioritize malfunction detection at the expense of classification accuracy.

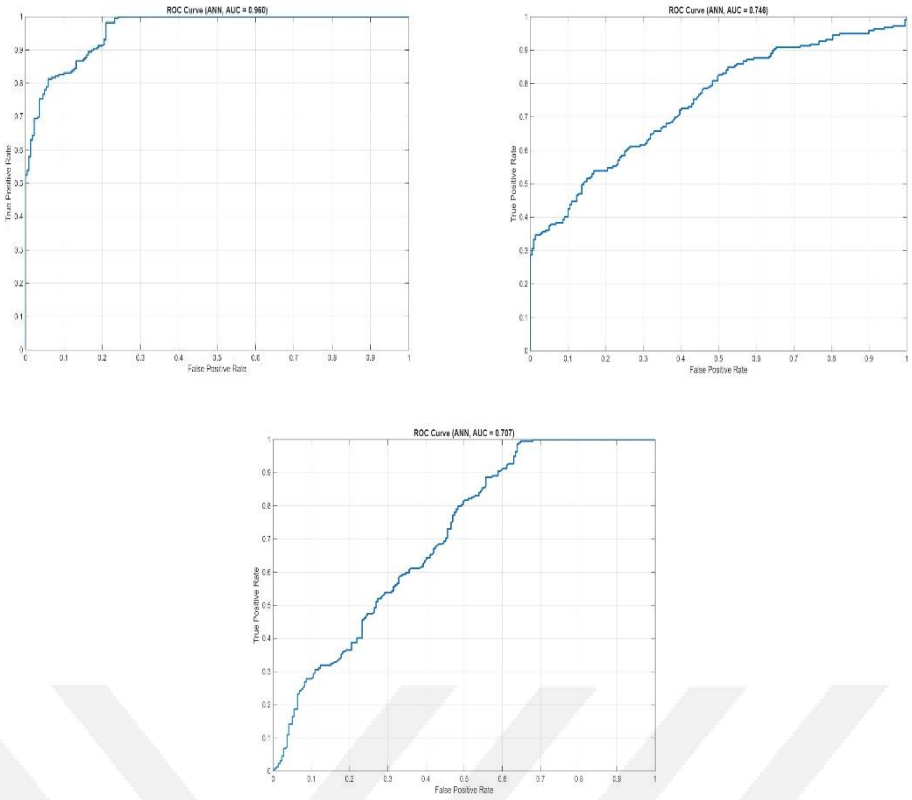


Figure 29 shows the ROC curves for the ANN model for the three days.

Figure 29 shows the ROC curves for the ANN model for the three days. The AUC = 0.960 on day 10, indicating excellent separation between the two cases. On day 56, the AUC dropped to 0.746, reflecting a decline in discrimination ability. On day 208, the AUC value was 0.707, indicating average performance, with relatively poor separation at some thresholds.

Table 8 Performance metrics of ANN anomaly detection model

Day	Accuracy	Precision	Recall	F1-score	AUC
10	0.8858	0.8238	0.9817	0.8958	0.9596
56	0.6598	0.6434	0.7169	0.6782	0.7461
208	0.6598	0.6012	0.9498	0.7363	0.7073

Table 8 summarizes the quantitative performance indicators of the ANN model. The results show that the model achieved the highest accuracy on day 10 with an accuracy of 0.8858 and an F1-score of 0.8958, due to the good balance between the

precision and recall values. On day 56, overall performance declined, with an accuracy of 0.6598 and a clear decrease in the F1-score, reflecting the instability of the model over time. On day 208, despite the high Recall value of 0.9498, the decrease in Precision led to an increase in false alarms, which may reduce the reliability of the model in practical applications.

Overall, the ANN model results show a clear improvement over the Autoencoder model, especially in detecting anomalies, but the fluctuation in performance across different days indicates the model's limited ability to generalize when signal characteristics change over time. This shortcoming highlights the need for more advanced models capable of representing complex temporal relationships, which will be addressed in the hybrid LSTM-CNN model in the next section.

5.4.4 Hybrid CNN–LSTM Anomaly Detection Model

The CNN–LSTM hybrid model combines two key features for processing complex dynamic system data, with convolutional neural networks (CNN) extract distinctive features from raw signals, while long-short term memory (LSTM) networks analyze temporal dependencies and sequential changes in system behavior. This combination enables the model to handle failures that do not appear as momentary outliers, but rather as cumulative temporal patterns that are difficult to detect using traditional models.

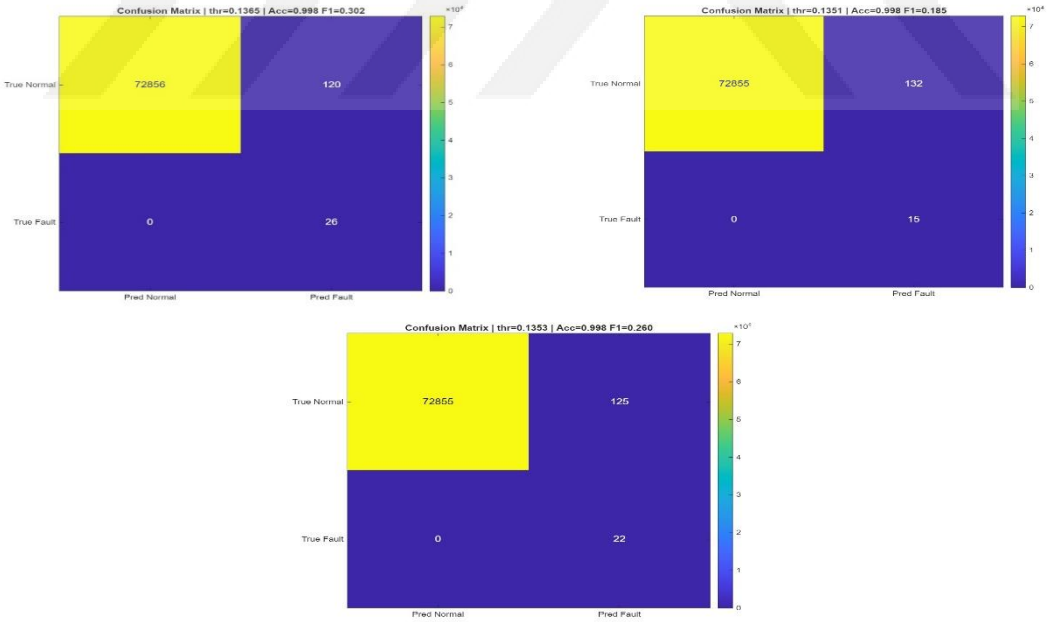


Figure 30 shows the confusion matrices for the Hybrid CNN–LSTM model for days (10, 56, 208)

The confusion matrix results show that the model achieved near-perfect detection of faults on all test days (10, 56, 208), with a Recall = 1.0 in all cases. This indicates that the model did not miss any actual malfunctions, which is a critical factor in critical system applications where ignoring a malfunction is more dangerous than triggering a false alarm.

On the other hand, a decrease in Precision values was observed, reflecting an increase in the number of false positives. However, this behavior is expected in early fault detection models, where high sensitivity is preferred over strict accuracy, especially in the early stages of fault development or in cases where the threshold signals are close to the fault region.

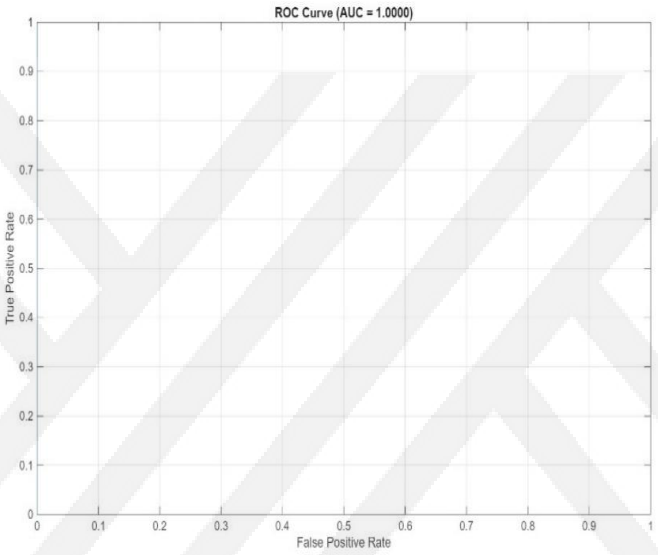


Figure 31 shows the ROC curves for the Hybrid CNN–LSTM model for prediction periods of 10, 56, and 208 days.

The AUC = 1.0 values across all days confirm the model's high ability to distinguish between the two cases in theory, indicating that the model learns a strong internal representation of the data structure. However, the combination of high AUC and low Precision indicates that the threshold used for decision-making has been set conservatively, with the aim of ensuring that no potential malfunction is missed.

Table 9 Performance metrics of Hybrid CNN–LSTM anomaly detection model

Day	Accuracy	Precision	Recall	F1-score	AUC
10	0.9984	0.1781	1.0000	0.3023	1.0000
56	0.9982	0.1020	1.0000	0.1852	1.0000
208	0.9983	0.1497	1.0000	0.2604	1.0000

Table 9 shows the quantitative performance indicators for the Hybrid CNN–LSTM model across the three test days (10, 56, 208). It can be seen that the model achieved very high accuracy values exceeding 99.8% on all days, which is due to the dominance of normal operating data compared to the number of fault cases, a common feature in real-world industrial system data. However, caution should be exercised when interpreting this high accuracy due to the unbalanced nature of the dataset. The relatively low accuracy values indicate the occurrence of some false alarms, although the model successfully detected all failure cases.

The Recall values of 1.0 show complete stability across all test days, indicating that the model successfully detected all malfunctions without exception. This behavior reflects the model's high sensitivity and ability to capture time patterns associated with malfunctions even in their early stages, which is a critical factor in preventive monitoring applications.

In contrast, the Precision values were relatively low, ranging between 0.10 and 0.18, indicating a high number of False Positives. However, this decrease does not reflect a weakness in the model so much as it reflects a conservative choice of decision threshold, aimed at reducing the likelihood of missing any actual failures at the expense of increasing warning alerts.

As a result of this balance, the F1-score values were moderate, affected by a decrease in Precision despite perfect Recall. This is expected behavior in early fault detection models, which are primarily designed to maximize sensitivity rather than strict accuracy.

On the other hand, AUC = 1.0 values on all days show the model's perfect ability to distinguish between the two cases from a probabilistic perspective, confirming that

the model has learned a strong internal representation of the data structure and that the apparent limitations in Precision are related to the final decision-making mechanism rather than a weakness in the model itself.

From an applied engineering perspective, the CNN–LSTM model can be considered particularly suitable for preventive monitoring systems, where the top priority is to detect all potential faults, even if this leads to an increase in false alarms. These results also open the door to improving the operational performance of the model by adjusting the threshold or integrating it with a subsequent classification model (Decision Layer) to reduce unnecessary alarms without affecting the basic detection capability.

Based on this, the CNN–LSTM model shows clear superiority in sensitivity and early detection compared to the other models used in this study, making it a strong choice for industrial applications that require the highest level of safety and rapid response to faults.

5.5 Overall Comparative Discussion and Final Remarks

The comparative Figure 32 showing the performance of fault detection algorithms on the test set reveals significant differences in each model's ability to distinguish between normal and fault conditions when evaluated using standard performance metrics: Accuracy, Precision, Recall, F1-score, and AUC.

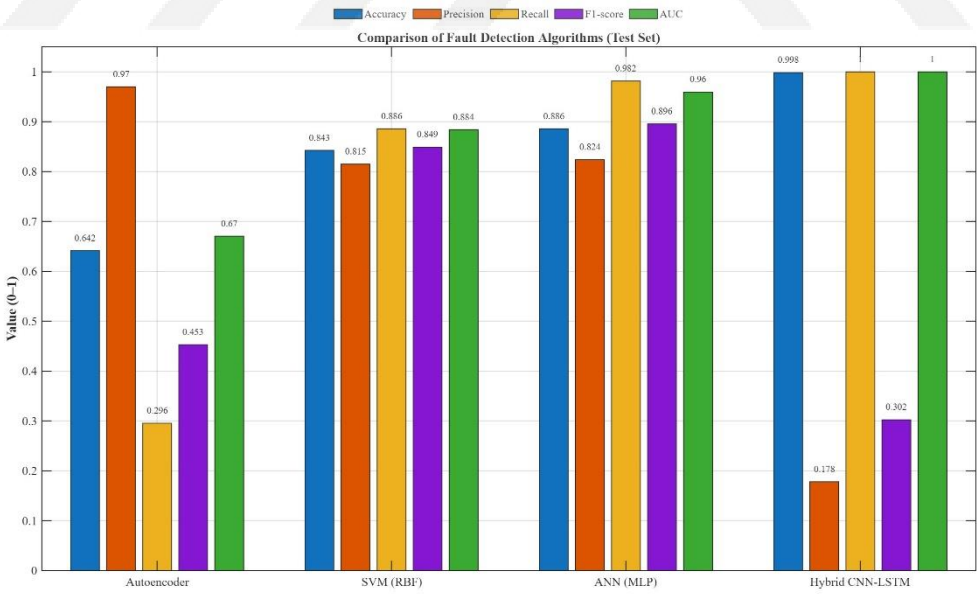


Figure 32 Comparison of fault detection algorithms on the test dataset

Table 10 Comparison of selected studies on wind turbine fault diagnosis.

Authors	Model	Data Type	Signals / Features	Accuracy	Key Contribution	Comparison with This Work
(Qi et al., 2024a)	CNN–LSTM	SCADA	Multi-sensor	90.06%	Multi-fault detection	The proposed model achieves higher accuracy (99.84%)
(Qin et al., 2023b)	CNN–LSTM (Parallel)	SCADA	Multi-sensor	99.60%	Parallel feature extraction	Comparable accuracy but without physics-based modeling
(Jiang, 2025)	CNN–LSTM	SCADA	Multi-sensor	96.5%	Hybrid deep learning model	The proposed model achieves higher accuracy and recall
(Nogueira et al., 2025)	Autoencoder	SCADA + Simulation	Multi-signal	99%	Early fault detection	Similar hybrid concept but different architecture
(Behara & Saha, 2025b)	CNN–LSTM + VMD	SCADA + Simulation	Electrical signals	88%	Real-time fault detection	Lower accuracy compared to proposed model
(S. Zhang et al., 2023)	T2V-LSTM	SCADA	Time-series	84.97%	Temporal encoding improves prediction	Lower performance compared to CNN-LSTM
(Alagha et al., 2026)	Multiscale CNN	Vibration	Time-frequency	F1 = 98.95%	Multiscale feature extraction	Comparable performance but different metric
(Y. Zhang et al., 2023)	LSTM-VAE-WGAN	SCADA	Multivariate	99.7% (confidence)	Handles noisy data	Uses anomaly threshold instead of classification
This Work	CNN–LSTM	Simulink	Temperature + Vibration	99.84%	Hybrid AI + physical modeling	—

As shown in Table 10 most previous studies achieved accuracies ranging from 84% to 96.5%, particularly those that relied on SCADA data and used deep learning models such as CNN–LSTM. For example, the study by (Qi et al., 2024b) achieved an

accuracy of 90.06%, while some advanced models such as parallel CNN–LSTM reached approxi-mately 99.6% (Qi et al., 2024b).

Compared to these studies, the model proposed in this research achieved an accu-racy of 99.84% along with a perfect recall value (1.000), reflecting a very high ability to detect all fault cases without missing any. This improvement is attributed to the inte-gration of physical modeling with deep learning techniques, which contributed to im-proving data quality and increasing the model's ability to represent the system's actual behavior.

However, it should be noted that the results of this study were based on simulation data generated using MATLAB/Simulink, whereas many of the reference studies relied on real SCADA data; therefore, direct comparisons between the results should be treated with caution.

Furthermore, most previous studies rely solely on SCADA data, whereas this work relies on data generated using a MATLAB/Simulink environment that includes thermal and vibration signals, providing a more comprehensive and accurate representation of various operating conditions. These results indicate that integrating physical modeling with deep learning hybrid models represents a promising direction for the development of fault diagnosis systems in wind turbines.

Compared to these studies, the model proposed in this research achieved perfor-mance among the highest recorded levels and even surpassed them in some cases, in-dicating that integrating physical modeling with deep learning provides an additional advantage over data-only methods.

Despite the high performance, the results revealed a challenge regarding the bal-ance between Recall and Precision, as the model tends to reduce Type II errors (failure to detect faults) at the expense of increasing Type I errors (false alarms) (Qi et al., 2024c). noted that this challenge is common in deep models and requires further improvements to achieve a better balance.

Compared to traditional methods, the proposed approach demonstrates clear su-periority due to its ability to learn automatically and detect complex, nonlinear pat-terns (Alagha et al., 2025b). Confirmed that hybrid models represent the most

advanced trend in the field of wind turbine fault diagnosis. However, there are some limitations in this study, as it relied on simulated data, which may not fully reflect the noise and complexity present in real data, a point con-sistent with what (Alagha et al., 2025b). Additionally, the number of failure cases studied was limited, which may affect the model's generalizability.

Accordingly, future studies are recommended to use real-world data from SCADA systems, as well as to develop techniques to improve the balance between Precision and Recall, such as using cost-sensitive learning algorithms or optimizing the decision threshold.

Overall, this study confirms that integrating physical modeling with hybrid deep learning models represents an effective and promising approach for improving fault diagnosis systems in wind turbines, and contributes to enhancing the reliability of re-newable energy systems and reducing maintenance costs.

CONCLUSIONS AND RECOMMENDATIONS

This thesis presented a comprehensive framework for diagnosing wind turbine failures by integrating physical modeling with hybrid deep learning techniques. A high-fidelity dynamic model was developed using the MATLAB/Simulink environment to simulate the thermal and mechanical behavior of the turbine under various operating conditions, enabling the generation of realistic data that contributed to the highly effi-cient training of AI models.

The results showed that the proposed hybrid model (CNN–LSTM) clearly outper-forms traditional models, achieving a high accuracy of 99.84% and a recall of 1.0000 across various time periods, demonstrating the model's ability to detect all failure cases without missing any. However, the results showed a relative decrease in precision, in-dicating the presence of some false alarms, which requires future improvement to achieve a better balance between sensitivity and precision.

Compared to previous studies, in which the accuracy of CNN–LSTM models typi-cally range between 90% and 96.5%, and in some advanced cases reaches approximately 99.6%, the model proposed in this study achieves competitive performance and even outperforms them in some cases. This improvement is attributed to the integration of physical modeling with deep learning techniques, which contributed to improving data quality and increasing the model's generalization ability.

From a practical standpoint, the proposed framework provides an effective tool for predictive maintenance systems, as it can detect early anomalies in temperature and vibration signals before they develop into serious failures, thereby helping to reduce unplanned outages and lower operating costs in wind farms.

Despite these positive results, the study faces some limitations, most notably the reliance on simulated data, which may not fully reflect the noise and complexity present in real SCADA system data. Additionally, the model was tested on a limited number of failure scenarios, which may affect its generalizability to more complex cases.

Therefore, the study recommends future work to validate the model's performance using real-world data, as well as developing techniques to improve the balance between Precision and Recall, such as using cost-sensitive learning or adjusting the decision threshold. Furthermore, integrating techniques such as transfer learning could enhance the model's ability to adapt to different operational environments.

Overall, this study confirms that integrating physical modeling with hybrid deep learning models represents an effective and practical approach to improving the diag-nosis of wind turbine faults and supports the development of smart maintenance sys-tems that contribute to enhancing the reliability of renewable energy systems.

REFERENCES

Al Lahham, E., Kanaan, L., Murad, Z., Khalid, H. M., Hussain, G. A., & Muyeen, S. M. (2025). Online condition monitoring and fault diagnosis in wind turbines: A comprehensive review on structure, failures, health monitoring techniques, and signal processing methods. *Green Technologies and Sustainability*, *3*(2), 100153. https://doi.org/10.1016/j.grets.2024.100153

Alagha, N., Khairuddin, A. S. M., Al-Khatib, O., & Copiaco, A. (2026). *CEEMDAN-Based Multiscale CNN for Wind Turbine Gearbox Fault Detection* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2601.06217

Alagha, N., Khairuddin, A. S. M., Haitaamar, Z. N., Al-Khatib, O., & Kanesan, J. (2025a). Artificial Intelligence in Wind Turbine Fault Detection and Diagnosis: Advances and Perspectives. *Energies*, *18*(7), 1680. https://doi.org/10.3390/en18071680

Allal, Z., Noura, H. N., Vernier, F., Salman, O., & Chahine, K. (2024). Wind turbine fault detection and identification using a two-tier machine learning framework. *Intelligent Systems with Applications*, *22*, 200372. https://doi.org/10.1016/j.iswa.2024.200372

An, X. L., Jiang, D. X., Li, S. H., & Chen, J. (2011). Fault diagnosis of direct-drive wind turbine based on support vector machine. *Journal of Physics: Conference Series*, *305*, 012030. https://doi.org/10.1088/1742-6596/305/1/012030

Aslam, S., Herodotou, H., Mohsin, S. M., Javaid, N., Ashraf, N., & Aslam, S. (2021). A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids. *Renewable and Sustainable Energy Reviews*, *144*, 110992. https://doi.org/10.1016/j.rser.2021.110992

Babu, G. V. R., Porlekar, S. B., Ali, H. M., Vuyyuru, V. A., Ansari, M. S. A., & Raj, I. I. (2024). Dynamic Fault Diagnosis in Wind Turbines: A GNN-LSTM Approach. *2024 IEEE 3rd International Conference on Electrical Power and Energy Systems (ICEPES)*, 1–6. https://doi.org/10.1109/ICEPES60647.2024.10653514

Behara, R. K., & Saha, A. K. (2025a). Comparative Performance Analysis of Deep Learning-Based Diagnostic and Predictive Models in Grid-Integrated Doubly Fed Induction Generator Wind Turbines. *Energies*, *18*(17), 4725. https://doi.org/10.3390/en18174725

Bindingsbø, O. T., Singh, M., Øvsthus, K., & Keprate, A. (2023). Fault detection of a wind turbine generator bearing using interpretable machine learning. *Frontiers in Energy Research*, *11*, 1284676. https://doi.org/10.3389/fenrg.2023.1284676

Binti Shahrulhisham, N. N. H., Chong, K. H., Yaw, C. T., & Koh, S. P. (2022). Application of Machine Learning Technique Using Support Vector Machine in Wind Turbine Fault Diagnosis. *Journal of Physics: Conference Series*, *2319*(1), 012017. https://doi.org/10.1088/1742-6596/2319/1/012017

Bouthaina Abichou, Flórez, D., Moamar Sayed-Mouchaweh, Houari Toubakh, François, B., & Girard, N. (2014). *Fault Diagnosis Methods for Wind Turbines Health Monitoring: A Review*. https://doi.org/10.13140/2.1.1150.6563

Chen, W., Qiu, Y., Feng, Y., Li, Y., & Kusiak, A. (2021). Diagnosis of wind turbine faults with transfer learning algorithms. *Renewable Energy*, *163*, 2053–2067. https://doi.org/10.1016/j.renene.2020.10.121

Chokr, B., Chatti, N., Charki, A., Lemenand, T., & Hammoud, M. (2024). Bi-LSTM Autoencoder SCADA based Unsupervised Anomaly Detection in Real Wind Farm Data. *2024 IEEE International Conference on Prognostics and Health Management (ICPHM)*, 174–183. https://doi.org/10.1109/ICPHM61352.2024.10626815

Craig, M. T., Cohen, S., Macknick, J., Draxl, C., Guerra, O. J., Sengupta, M., Haupt, S. E., Hodge, B.-M., & Brancucci, C. (2018). A review of the potential impacts of climate change on bulk power system planning and operations in the United States. *Renewable and Sustainable Energy Reviews*, *98*, 255–267. https://doi.org/10.1016/j.rser.2018.09.022

Dallabona, A., Blanke, M., Pedersen, H. C., & Papageorgiou, D. (2025). Fault diagnosis and prognosis capabilities for wind turbine hydraulic pitch systems. *Mechanical Systems and Signal Processing*, *224*, 111941. https://doi.org/10.1016/j.ymssp.2024.111941

Dao, P. B. (2023). On Cointegration Analysis for Condition Monitoring and Fault Detection of Wind Turbines Using SCADA Data. *Energies*, *16*(5), 2352. https://doi.org/10.3390/en16052352

Eriksson, S., Bernhoff, H., & Leijon, M. (2008). Evaluation of different turbine concepts for wind power. *Renewable and Sustainable Energy Reviews*, *12*(5), 1419–1434. https://doi.org/10.1016/j.rser.2006.05.017

Fadzail, N. F., Zali, S. M., Mid, E. C., & Bakar, N. N. A. (2025). Hyperparameter optimization using automated machine learning based artificial neural network for fault detection in wind turbine generator. *Next Research*, *2*(4), 100785. https://doi.org/10.1016/j.nexres.2025.100785

García-Vaca, M. A., Sierra-García, J. E., & Santos, M. (2026). Probabilistic evaluation for early wind turbine yaw misalignment detection. *Reliability Engineering & System Safety*, *266*, 111716. https://doi.org/10.1016/j.ress.2025.111716

Ghanbari, N., Riyazi, Y., Shirazi, F. A., & Kalhor, A. (2023). *Enhanced Gearbox Fault Diagnosis with Fusion LSTM-CNN Network.*

Jiang, Q. (2025). Hybrid CNN-LSTM Model for Predictive Maintenance of Wind Turbine Systems. *International Journal of Engineering, Science and Information Technology*, *5*(4), 517–526. https://doi.org/10.52088/ijesty.v5i4.1679

Laouti, N., Sheibat-Othman, N., & Othman, S. (2011). Support Vector Machines for Fault Detection in Wind Turbines. *IFAC Proceedings Volumes*, *44*(1), 7067–7072. https://doi.org/10.3182/20110828-6-IT-1002.02560

Liu, J., Wang, X., Wu, S., Wan, L., & Xie, F. (2023). Wind turbine fault detection based on deep residual networks. *Expert Systems with Applications*, *213*, 119102. https://doi.org/10.1016/j.eswa.2022.119102

Liu, J.-H., Corbita, N. T., Lee, R.-M., & Wang, C.-C. (2022). Wind Turbine Anomaly Detection Using Mahalanobis Distance and SCADA Alarm Data. *Applied Sciences*, *12*(17), 8661. https://doi.org/10.3390/app12178661

Malik, Dr. P. (2024). Wind Turbine Failure Prediction Using Time-Series Analysis and Deep Learning: A Predictive Maintenance Approach. *International Journal of Science, Engineering and Technology*, *12*(6), 1–14. https://doi.org/10.61463/ijset.vol.12.issue6.403

Morrison, R., Liu, X., & Lin, Z. (2022). Anomaly detection in wind turbine SCADA data for power curve cleaning. *Renewable Energy*, *184*, 473–486. https://doi.org/10.1016/j.renene.2021.11.118

Mulongo, J., Atemkeng, M., Ansah-Narh, T., Rockefeller, R., Nguegnang, G. M., & Garuti, M. A. (2020). Anomaly Detection in Power Generation Plants Using Machine Learning and Neural Networks. *Applied Artificial Intelligence*, *34*(1), 64–79. https://doi.org/10.1080/08839514.2019.1691839

Nabi, K. N., Tahmid, M. T., Rafi, A., Kader, M. E., & Haider, Md. A. (2021). Forecasting COVID-19 cases: A comparative analysis between recurrent and convolutional neural networks. *Results in Physics*, *24*, 104137. https://doi.org/10.1016/j.rinp.2021.104137

Ng, E. Y.-K., & Lim, J. T. (2022). Machine Learning on Fault Diagnosis in Wind Turbines. *Fluids*, *7*(12), 371. https://doi.org/10.3390/fluids7120371

Nogueira, W. F., Melani, A. H. D. A., & De Souza, G. F. M. (2025). Wind Turbine Fault Detection Through Autoencoder-Based Neural Network and FMSA. *Sensors*, *25*(14), 4499. https://doi.org/10.3390/s25144499

Ohunakin, O. S., Henry, E. U., Matthew, O. J., Ezekiel, V. U., Adelekan, D. S., & Oyeniran, A. T. (2024). Conditional monitoring and fault detection of wind turbines based on Kolmogorov–Smirnov non-parametric test. *Energy Reports*, *11*, 2577–2591. https://doi.org/10.1016/j.egyr.2024.01.081

Pereira, J., & Silveira, M. (2018). Unsupervised Anomaly Detection in Energy Time Series Data Using Variational Recurrent Autoencoders with Attention. *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 1275–1282. https://doi.org/10.1109/ICMLA.2018.00207

Pozo, F., & Vidal, Y. (2015). Wind Turbine Fault Detection through Principal Component Analysis and Statistical Hypothesis Testing. *Energies*, *9*(1), 3. https://doi.org/10.3390/en9010003

Puruncajas, B., Castellani, F., Vidal, Y., & Tutivén, C. (2025). Use of Artificial Neural Networks and SCADA Data for Early Detection of Wind Turbine Gearbox Failures. *Machines*, *13*(8), 746. https://doi.org/10.3390/machines13080746

Qi, L., Zhang, Q., Xie, Y., Zhang, J., & Ke, J. (2024a). Research on Wind Turbine Fault Detection Based on CNN-LSTM. *Energies*, *17*(17), 4497. https://doi.org/10.3390/en17174497

Qin, S., Tao, J., & Zhao, Z. (2023a). Fault diagnosis of wind turbine pitch system based on LSTM with multi-channel attention mechanism. *Energy Reports*, *10*, 4087–4096. https://doi.org/10.1016/j.egyr.2023.10.076

Rama, V. S. B., Hur, S.-H., & Yang, J.-M. (2024). Short-Term Fault Prediction of Wind Turbines Based on Integrated RNN-LSTM. *IEEE Access*, *12*, 22465–22478. https://doi.org/10.1109/ACCESS.2024.3364395

Reddy, A., Begum, S., Juntupally, S., & Anupoju, G. R. (2024). Assessing the impact of alkali pretreatment of rice husk on its composition and product portfolio: Tradeoff between biogas and valuable materials recovery for sustainability. *Renewable Energy*, *226*, 120393. https://doi.org/10.1016/j.renene.2024.120393

Rossi, B., Chren, S., Buhnova, B., & Pitner, T. (2016). Anomaly detection in Smart Grid data: An experience report. *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 002313–002318. https://doi.org/10.1109/SMC.2016.7844583

Rudin, K., Hua, M.-D., Ducard, G., & Bouabdallah, S. (2011). A Robust Attitude Controller and its Application to Quadrotor Helicopters. *IFAC Proceedings Volumes*, *44*(1), 10379–10384. https://doi.org/10.3182/20110828-6-IT-1002.02377

Sheng, L., Li, C., Gao, M., Xi, X., & Zhou, D. (2025). A review of SCADA-based condition monitoring for wind turbines via artificial neural networks. *Neurocomputing*, *633*, 129830. https://doi.org/10.1016/j.neucom.2025.129830

Stetco, A., Dinmohammadi, F., Zhao, X., Robu, V., Flynn, D., Barnes, M., Keane, J., & Nenadic, G. (2019). Machine learning methods for wind turbine condition monitoring: A review. *Renewable Energy*, *133*, 620–635. https://doi.org/10.1016/j.renene.2018.10.047

Villanueva, D., & Feijóo, A. (2020). A Review on Wind Turbine Deterministic Power Curve Models. *Applied Sciences*, *10*(12), 4186. https://doi.org/10.3390/app10124186

Wadawa, B., & Effa, J. Y. (2025). New strategy for fault detection and classification in wind turbines based on doubly-fed induction generators. *Global Energy Interconnection*, *8*(4), 668–684. https://doi.org/10.1016/j.gloei.2025.04.004

Wang, A., Pei, Y., Zhu, Y., & Qian, Z. (2023). Wind turbine fault detection and identification through self-attention-based mechanism embedded with a

multivariable query pattern. *Renewable Energy*, *211*, 918–937. https://doi.org/10.1016/j.renene.2023.05.003

Wang, B., Zhou, B., Zhu, D., Zou, M., Rao, Z., & Luo, H. (2025). Feature alignment LSTM for detecting anomalies in wind turbine main bearing temperature based on SCADA data. *Journal of Engineering Research*, S2307187725000628. https://doi.org/10.1016/j.jer.2025.05.011

Xiang, L., Wang, P., Yang, X., Hu, A., & Su, H. (2021). Fault detection of wind turbine based on SCADA data analysis using CNN and LSTM with attention mechanism. *Measurement*, *175*, 109094. https://doi.org/10.1016/j.measurement.2021.109094

Yan, J., Liu, Y., Meng, H., Li, L., & Ren, X. (2024). Wind turbine generator early fault diagnosis using LSTM-based stacked denoising autoencoder network and stacking algorithm. *International Journal of Green Energy*, *21*(11), 2477–2492. https://doi.org/10.1080/15435075.2024.2315445

Yin, A., Yan, Y., Zhang, Z., Li, C., & Sánchez, R.-V. (2020). Fault Diagnosis of Wind Turbine Gearbox Based on the Optimized LSTM Neural Network with Cosine Loss. *Sensors*, *20*(8), 2339. https://doi.org/10.3390/s20082339

Zare, S., & Ayati, M. (2021). Simultaneous fault diagnosis of wind turbine using multichannel convolutional neural networks. *ISA Transactions*, *108*, 230–239. https://doi.org/10.1016/j.isatra.2020.08.021

Zhang, C., & Yang, T. (2023). Anomaly Detection for Wind Turbines Using Long Short-Term Memory-Based Variational Autoencoder Wasserstein Generation Adversarial Network under Semi-Supervised Training. *Energies*, *16*(19), 7008. https://doi.org/10.3390/en16197008

Zhang, S., Robinson, E., & Basu, M. (2023). Wind Turbine Predictive Fault Diagnostics Based on a Novel Long Short-Term Memory Model. *Algorithms*, *16*(12), 546. https://doi.org/10.3390/a16120546

Zhang, X., Han, P., Xu, L., Zhang, F., Wang, Y., & Gao, L. (2020). Research on Bearing Fault Diagnosis of Wind Turbine Gearbox Based on 1DCNN-PSO-SVM. *IEEE Access*, *8*, 192248–192258. https://doi.org/10.1109/ACCESS.2020.3032719

Zhang, Y., Huang, R., & Li, Z. (2023). Fault Detection Method for Wind Turbine Generators Based on Attention-Based Modeling. *Applied Sciences*, *13*(16), 9276. https://doi.org/10.3390/app13169276

Zhang, Y., Liu, W., Wang, X., & Gu, H. (2022). A novel wind turbine fault diagnosis method based on compressed sensing and DTL-CNN. *Renewable Energy*, *194*, 249–258. https://doi.org/10.1016/j.renene.2022.05.085

RESUME

Personal Information

Surname, name : Zeyad Tareq Ahmed Ahmed

Nationality : Iraq

Birth date and place :

Telephone :

Fax :

E-mail :

Education

Degree	Education Unit	Graduation Date
Master	Istanbul Gelisim University	2026
Bachelor	Kirkuk Technical Engineering College	2021
High School	Private schools	2017

Work Experience

Year	Place	Title
2016\2019	Cell Phone Repair Technician and Programmer	Technical Engineer
2019/2021	Programming and Installation of SCADA Systems in Oil and Gas Refineries	Team Leader
2021\2023	Deputy Director of the Kirkuk Oil Institute	Team Leader

Foreign Language

I am a native Kurdish speaker and fluent in Arabic. I have a B2 level of proficiency in English and basic communication skills in Turkish.

Publications

The research conducted in this thesis has resulted in the publication of the following peer-reviewed journal article:

Ahmed, Z. T., & Aykut, E. (2026). Hybrid Deep Learning-Based Fault Detection in Wind Turbines Using Simulink-Generated Data. Applied Sciences, 16(11), 5438. https://doi.org/10.3390/app16115438

Hobbies

Technology and innovation, traveling, video gaming, and exploring new developments in science and technology and learn more language.