



OPEN Intelligent fault prediction and diagnosis for wind-powered heating systems using graph neural networks

Yuechao Wang^{1✉}, Jizhong Zhao^{1✉}, Donglai Tang², Weijie Zhao³ & Shan Huang⁴

With the rapid global transition towards clean energy, wind-powered heating systems have emerged as a critical solution for efficient wind energy utilization, particularly in the northern regions of China. However, these systems face significant reliability challenges due to complex spatiotemporal couplings and harsh operating conditions. This paper presents an adaptive fault prediction and intelligent diagnosis method based on a Multi-level Spatiotemporal Graph Neural Network to address the challenges of multi-source data fusion difficulties and inadequate spatiotemporal feature extraction. The proposed framework establishes a dynamic adaptive threshold generation mechanism by integrating maximum a posteriori probability estimation with interquartile range analysis, enabling real-time system state monitoring and early fault warning. The methodology incorporates graph attention networks, seven-branch parallel subgraph architectures, and multi-head attention mechanisms to capture topological evolution patterns through dynamic graph neural networks, while temporal attention modules are employed to enhance sequential dependencies of critical parameters. Experimental validation was conducted using 42 TB of SCADA data from China Guoneng Group's 200 MW wind-heat cogeneration project. The results demonstrate the model's superior multi-level diagnostic capability, achieving a comprehensive prediction accuracy of 93.5% and a fault detection F β -score ($\beta = 0.5$) of 0.95—representing an 18.6% improvement over traditional approaches—while maintaining strong robustness (KL divergence 0.09 ± 0.02) under transient operating conditions.

Keywords Wind-heat integrated energy systems, Graph neural networks (GNN), Fault prediction and diagnosis, Dynamic adaptive threshold, Spatiotemporal feature fusion

Background and Motivation

With the accelerated transformation of the global energy structure towards cleanliness, wind power heating systems, as an important form of efficient utilization of wind energy, have been widely used in the “Three North” regions of China¹. However, the coupling system between wind turbines and heating pipelines under complex climate conditions faces severe reliability challenges: on the one hand, extreme temperature differences ($-30\text{ }^{\circ}\text{C}$ – $50\text{ }^{\circ}\text{C}$) and frequent start stop conditions lead to a mechanical component failure rate that is more than 40% higher than that of conventional wind farms². On the other hand, the strong coupling characteristics between pressure fluctuations in the heating network and the operating status of fans result in a high false alarm rate of up to 35% for traditional single equipment monitoring methods³. The current operation and maintenance mode mainly relies on regular maintenance and post failure maintenance, but the International Energy Agency report shows that this passive maintenance makes the operation and maintenance cost of wind and heat systems account for 28% of power generation revenue, far exceeding that of new energy systems such as photovoltaics^{4,5}.

Related works

Currently, data-driven methods have become a popular technology in fault diagnosis for wind power heating systems. Based on data analysis techniques, these methods deeply mine offline and online big data to identify fault characteristics and determine the cause, location, and time of faults. This approach is generally categorized

¹College of Computer Science and Technology, Xi'an Jiaotong University, Xi'an 710049, Shaanxi, China. ²Aostar Information Technology Co., Ltd., Chengdu 610299, Sichuan, China. ³School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, Shaanxi, China. ⁴Software Engineering, Shanghai Jiaotong University, Shanghai 200240, China. ✉email: xjtu_xiaqi@163.com; gsfgwer112@163.com

into two major types: signal processing methods and big data methods^{6,7}. Signal processing-based methods analyze signals captured by monitoring equipment, such as lubrication conditions, acoustic emissions, and vibration signals, to diagnose faults by comparing signal characteristics under normal and faulty conditions^{8,9}. For instance, Yang et. al.¹⁰ proposed a novel sparse time–frequency representation method based on sparse representation theory and Wigner Ville distribution. By constructing a joint redundant dictionary and sparse decomposition technique, the early fault diagnosis accuracy of wind turbines was effectively improved. Simulation and field experiments verified the superiority and robustness of this method in engineering applications. Kong et. al.¹¹ proposed an enhanced sparse representation intelligent recognition method, which effectively solves the problem of fault diagnosis of planetary bearings in wind turbine gearboxes through structured dictionary design and sparse reconstruction error minimization discrimination criteria. However, signal processing methods are still subject to various factors such as high cost, high durability of sensors, and insufficient detection signals.

The diagnostic methods based on big data technology mainly rely on statistical, machine learning, deep learning and other technologies to highly automate the analysis of massive data, transforming the data obtained from state detection into relevant models reflecting degradation behavior. Due to the lack of installation of additional data collection equipment, data-driven big data methods based on data collection and monitoring systems have become a relatively economical and applicable fault diagnosis method for wind turbines^{12,4}. Liang et. Al.¹³ proposed a fault diagnosis algorithm for wind power converters based on Dempster Shafer evidence theory and Deng entropy fusion multi-scale approximate entropy. By mining multi-scale signal features, fusing conflict information, and optimizing weight allocation, it demonstrates higher diagnostic accuracy and stronger robustness than existing methods in simulation and experimental data. Hsu et al.¹⁴ integrated statistical process control with machine learning techniques to develop a fault diagnosis and maintenance prediction model by analyzing 2.8 million sensor data points from 31 wind turbines in Taiwan from 2015 to 2017, achieving an accuracy rate of 92.68% and enabling early warning of wind turbine abnormalities and improved maintenance efficiency. Xiao et. Al.¹⁵ proposed an AOC ResNet50 network model based on an improved attention octal convolution structure (AOctConv), which detects inverter faults by processing radar images generated from wind turbine SCADA system data. The detection accuracy reaches 98%, effectively solving the problem of information asymmetry in traditional methods for feature extraction in high and low frequency domains. Although manual annotation can alleviate the issue of scarce fault labels, the complexity of fault data scenarios requires substantial time investment and high professional expertise from technical personnel. Therefore, developing lower-cost methods that require no manual intervention represents the future direction in this field¹⁶.

In recent years, Normal Behavior Modeling (NBM) has gained increasing attention as a fault detection method that utilizes only historical healthy data to establish models. The NBM approach typically consists of two phases: prediction and fault detection^{17,18}. During the prediction phase, NBM constructs a normal behavior prediction model based on historical healthy data and calculates residuals between predicted and measured values. In the fault detection phase, NBM builds a normal distribution model using these residuals to identify outliers, thereby determining whether the target wind turbine is in an abnormal state. Consequently, NBM-based fault detection is regarded as a promising low-cost solution. Reference¹⁹ proposed a method combining correlation vector machine and residual based wind turbine fault detection and isolation. Reference²⁰ proposed a method of using stacked autoencoders to identify health status and diagnose faults in rotating mechanical components.

Contributions

Despite these advancements, several critical research gaps remain unresolved, particularly for wind-powered heating systems. Firstly, although recent work has made progress in graph based fault diagnosis, they mainly focus on individual components and overlook the dynamic thermoelectric coupling effects between wind turbines and heating networks²¹. Secondly, existing graph neural network methods typically rely on static graph structures²² or learning graphs without physical constraints, which cannot capture time-varying correlations under fluctuating operating conditions²³. Thirdly, most threshold methods are still static or semi adaptive²⁴, lacking online learning mechanisms to automatically adapt to the non-stationary behavior of the system. Finally, although some methods address temporal or spatial dependencies, few effectively combine multi-scale spatiotemporal feature learning with physical interpretability²⁵. To address the critical challenges of multi-source data fusion difficulties and inadequate spatiotemporal feature extraction in fault diagnosis for wind-powered heating systems, this paper innovatively proposes an intelligent diagnostic method based on multi-level spatiotemporal graph neural networks, with core contributions manifested in two aspects:

- (1) Unlike general spatiotemporal GNNs that typically rely on learning or distance based graphs, this paper constructs a hierarchical graph structure based on the actual physical topology of wind power heating systems. This structure clearly models the interactions between device level, subsystem level, and global system level.
- (2) To the limitations of traditional data-driven fault diagnosis methods in using static thresholds for anomaly detection, this paper proposes a dynamic adaptive threshold generation mechanism based on the fusion of Bayesian maximum a posteriori probability estimation and quartile range analysis. This mechanism can achieve online autonomous calibration of threshold parameters based on real-time system operating status and historical health data, effectively overcoming the inherent limitation of static threshold strategies in adapting to highly variable operating conditions of wind power heating systems.

Description of the online fault detection framework for wind power heating systems

In this paper, a graph neural network model based on GAT mechanism, seven parallel subgraph networks and multi-head attention mechanism is proposed to capture the topological dynamic characteristics of wind power

heating system through dynamic graph neural network and time series attention module, and generate adaptive thresholds by combining the a priori probability model with the online learning strategy to realise the intelligent fault diagnosis and early warning of the system's multi-source data as shown in Fig. 1.

Predictive modelling of normal behaviour

In this paper, a spatio-temporal graphical neural network model of wind power heating system is constructed to predict the critical state quantities (including unit power, bearing temperature, pipe network pressure, etc.) in the next time window using the multivariate parameters of the units and thermal equipment in the historical time window as inputs to realise the system-level fault prediction. The model can be described as:

$$\hat{Y} = f(v, A) \tag{1}$$

where: $v = [x_1, x_2, \dots, x_i, \dots, x_N]$ denotes the input node feature matrix, N is the total number of sensors in the SCADA system, and x_i is the node feature of sensor i ; A is the adjacency matrix; $\hat{Y} = [p_t, p_{t+1}, \dots, p_{t+T}]$ denotes the output label matrix of the prediction model, where p_{t+T} is the predicted value of the state parameters of the WTGs in the moment of $t+T$, and T is the time window size.

Adaptive threshold generation method for wind power heating anomalies

This paper presents an innovative fault detection approach for wind-powered heating systems, which establishes component-level normal behavior models through a novel integration of Bayesian maximum a posteriori estimation, multi-scale spatiotemporal graph neural networks, and historical operational data. The methodology dynamically determines adaptive detection thresholds using the interquartile range (IQR) technique, enabling automatic threshold calibration in response to varying operational conditions. This framework facilitates comprehensive real-time condition monitoring and early fault detection across the entire system infrastructure, encompassing both wind turbine generators and thermal energy delivery components.

A priori probability model construction

Firstly, a Z-score model was applied to the historical health data for standardisation, and then a normal distribution model as shown in Eq. (2) was built based on these data.

$$p(y') = \frac{\exp\left(-\frac{(y' - \mu)^2}{2\sigma^2}\right)}{\sqrt{2\pi}\sigma} \tag{2}$$

where μ and σ are the mean and standard deviation of the normal historical data, respectively, y' is the predicted value of the state parameter under the normal state of the wind power heating system, $p(y')$ is the probability distribution function of y' .

Construction of normal distribution model for output power of wind power heating system under normal state

This paper develops a probabilistic modeling approach for wind-powered heating systems by combining multi-scale spatiotemporal graph neural networks with operational history. Under the normality assumption of healthy system data, we implement a dynamic parameter estimation framework using maximum a posteriori probability to continuously update the median values of component state distributions. The methodology

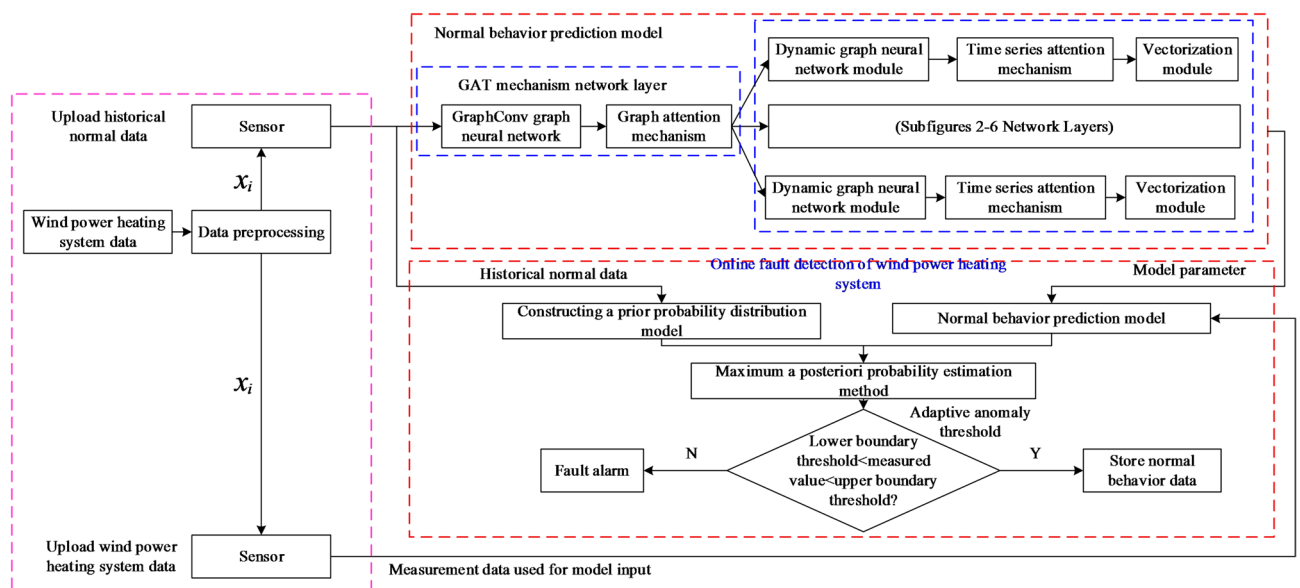


Fig. 1. Overall framework of online fault detection of wind turbines.

effectively integrates both historical operational patterns and real-time predictive analytics for comprehensive system health assessment.

$$\begin{aligned}\hat{\theta}_{MAP} &= \arg \max_{\theta} (p(\hat{Y}|\theta) p(\theta)) \\ &= \arg \max_{\theta} \left(\log \prod_{P_t} \Pr(P_t|\theta) \Pr(\theta) \right) \\ &= \arg \max_{\theta} \left(\sum_{P_t} (\log \Pr(P_t|\theta)) + \log \Pr(\theta) \right)\end{aligned}\quad (3)$$

where $p(\hat{Y}|\theta)$ denotes the probability density of \hat{Y} under the normal state of the wind power heating system with the given model parameter θ ; $\hat{\theta}_{MAP}$ is the median of the normal distribution of \hat{Y} under the normal state of the wind power heating system; P_t is the predicted state parameter under the normal state of the wind power heating system at time t ; and $\Pr(P_t|\theta)$ denotes the probability of the predicted value of P_t to occur with the given model parameter θ ; $\Pr(\theta)$ denotes the prior probability distribution under a given model parameter θ .

Adaptive threshold generation for abnormal states in wind power heating systems

Based on the constructed normal distribution model of the state parameters of the wind power heating system, the IQR method is used to dynamically generate the anomaly detection thresholds, which effectively improves the accuracy of the online fault detection of the system by integrating the historical operation experience with the real-time prediction data.

$$CDF(P_t) = \frac{1}{\sqrt{2\pi}\sigma_t} \int_{-\infty}^{P_t} \exp\left(-\frac{(P_t - \mu_t)^2}{2\sigma_t^2}\right) dP_t \quad (4)$$

$$\begin{cases} CDF_{P_t}(Q_{3,t}) - CDF_{P_t}(Q_{1,t}) = 50\% \\ Q_{3,t} - \mu_t = \mu_t - Q_{1,t} \end{cases} \quad (5)$$

$$\begin{cases} V_{UT,t} = Q_{3,t} + 1.5\delta_{IQR} \\ V_{LT,t} = Q_{1,t} - 1.5\delta_{IQR} \end{cases} \quad (6)$$

where $CDF(\cdot)$ represents the cumulative probability density function; $CDF_{P_t}(\cdot)$ represents the probability that the value of the state parameter P_t at time t is less than or equal to a specific state parameter; σ_t is the standard deviation of the state parameters at time t ; $Q_{1,t}$ is the upper quartile value at time t ; $Q_{3,t}$ is the lower quartile value at time t ; μ_t is the median of the normal distribution model of the target state parameters of the wind turbine at time t ; δ_{IQR} is the difference between $Q_{1,t}$ and $Q_{3,t}$; $V_{UT,t}$ and $V_{LT,t}$ are the upper and lower boundary thresholds of the wind power heating system, respectively.

Fault prediction based on multi-layer spatio-temporal graph neural network

In this paper, a multi-level graph neural network model based on the physical structure of the wind power heating system is constructed. By dividing the 37 key parameters monitored by the SCADA system (including pitch, nacelle, gearbox, and other unit data) into three levels: global sensor interconnection diagram, functional unit sub-diagram, and global interconnection feature diagram, we achieve a high degree of coupling between the system's physical structure and the prediction model, and effectively mine the spatial and temporal correlation characteristics of the wind turbine operation data in the heating system.

GAT mechanism network layer

The normal behaviour prediction model for wind power heating system proposed in this paper uses a dedicated GAT mechanism network layer for high-density data optimisation, which integrates the GraphConv graph neural network module and the GAT module supported by the Masked Attention mechanism, in which the GraphConv module optimises the self-node information transfer and the neighbourhood node information aggregation through two independently learnable weight parameters W_1 and W_2 , respectively. Neighbourhood node information aggregation, compared with the traditional GCN model can be more effective in dealing with the complex topological relationships between the functional units in the wind power heating system, thus improving the model's ability to extract spatio-temporal features of the system.

$$s_i = \sigma \left(W_1 v_i + W_2 \sum_{j \in N_i} e_{j,i} v_j \right) \quad (7)$$

where $\sigma(\cdot)$ denotes the activation function; v_j is the transmission feature of the hidden layer of the neighbour node j ; v_i is the transmission feature of the hidden layer of the self node i ; $e_{j,i}$ denotes the connectivity state between nodes j and i . To find the degree of influence between any two sensor nodes, all the $e_{j,i}$ will be regarded as 1; s_i denotes the reconstructed feature of the hidden layer of the node i ; and N_i denotes the set of the

neighbouring nodes of node i . By using a loss function such as the mean square error, the model captures the strength of the interaction between each functional unit and sensor node more accurately.

$$s_{i'} = \sigma \left(\frac{1}{K} \sum_{k=1}^K \sum_{j \in N_i} \alpha_{ij,k} W_k s_{i,j} \right) \tag{8}$$

where $s_{i,j}$ denotes the reconstructed features of the hidden layer of node j , the neighbour of node i ; $a^T (W s_i // W s_{i,j})$ denotes the strength of association between these two nodes in the GAT mechanism by splicing the feature vectors of nodes i and j transformed by the weight matrix W and dot-producting them with a learnable weight vector a ; K denotes the number of attention heads in the GAT mechanism; $s_{i'}$ denotes the final output features of node i in the network layer of the GAT mechanism; $\alpha_{ij,k}$ denotes the attention coefficient between nodes i and j in the attention head k ; W_k is the trainable weight parameter of attention head k .

Dynamic graph neural network model

In the updating memory phase, the reconstructed features of spatio-temporal features within the same subgraph are obtained by keeping the information that was selectively forgotten in the previous moment linked to the information that is selectively remembered in the current moment.

$$X_{d,t'} = \sigma \left(\tilde{D}_{d,t}^{-\frac{1}{2}} \hat{A}_{d,t} \tilde{D}_{d,t}^{-\frac{1}{2}} X_{d,t} \theta_d \right) \tag{9}$$

$$z_t = \sigma (W_z [X_{d,t'}, h_{d,t-1}]) \tag{10}$$

$$r_t = \sigma (W_r [X_{d,t'}, h_{d,t-1}]) \tag{11}$$

$$\tilde{h}_t = \tanh (W_{\tilde{h}} [r_t h_{d,t-1}, X_{d,t'}]) \tag{12}$$

$$h_{d,t} = (1 - z_t) h_{d,t-1} + z_t \tilde{h}_t \tag{13}$$

where: $\hat{A}_{d,t}$ is the adjacency matrix of subgraph d at time t ; $\tilde{D}_{d,t}$ is the degree of subgraph d at time t ; $\tilde{D}_{d,t}^{-\frac{1}{2}} \hat{A}_{d,t} \tilde{D}_{d,t}^{-\frac{1}{2}}$ denotes the edge-based normalisation operation; $X_{d,t}$ is the losing feature of subgraph d at time t ; $X_{d,t'}$ is the feature of subgraph d at time t after undergoing the reconstruction of the GCNConv model; θ_d is the training parameter of subgraph d in the GCNConv model. At time step t , \tilde{h}_t represents the potential hidden state, while W_z , W_r , and $W_{\tilde{h}}$ correspond to the weight matrices for the update gate, reset gate, and candidate hidden state, respectively.

Time series attention mechanism

By introducing a time series attention mechanism module, the temporal characteristics of each time step are further strengthened, enhancing the correlation of features at each training time step. Firstly, the time window features of subgraph d are mapped into a parameter matrix ed through a two-layer network. After normalization, the temporal weight parameters are dynamically fused into the original input features, effectively enhancing the temporal correlation representation ability of wind power heating system operation data.

$$e_d = w_{(2)} (w_{(1)} H_d + b_{(1)}) + b_{(2)} \tag{14}$$

$$\alpha_{d,t} = \frac{\exp(e_{d,t})}{\sum_{t=1}^T \exp(e_{d,t})} \tag{15}$$

$$C_{d,t} = \sum_{t=1}^T \alpha_{d,t} h_{d,t} \tag{16}$$

where $H_d = \{h_{d,t}, h_{d,t+1}, \dots, h_{d,t+T}\}$ is the hidden-layer input state features of subgraph d ; $e_d = \{e_{d,t}, e_{d,t+1}, \dots, e_{d,t+T}\}$ is the matrix of the hidden-layer mapped state features of subgraph d . $w_{(1)}$ and $b_{(1)}$ are the weight parameter and the bias of the network mapping of layer 1, respectively. $w_{(2)}$ and $b_{(2)}$ are the weight parameter and the bias of the network mapping of layer 2, respectively; $\alpha_{d,t}$ is the weight of the features of subgraph d in the whole time window at time t ; $C_{d,t}$ is the time window feature of the embedded correlation feature of subgraph d .

In this article, the MLS-GNN architecture is chosen to solve the problem of fault diagnosis in wind power heating systems for the following reasons:

Graph structure and physical topology: A multi-level graph is a direct computational representation of the physical hierarchical structure of a wind power heating system. Nodes correspond to physical sensors, while edges represent functional coupling or physical proximity.

Dynamic graph convolution and adaptive coupling: The dynamic characteristics of adjacency matrix A (t) aim to capture the time-varying coupling strength between the wind field and the heating network. Dynamic graph convolution automatically adapts to these changes and effectively learns non-stationary, state dependent physical relationships between components.

GAT mechanism and fault propagation: The attention weight α_{ij} learned in the GAT layer quantifies the dynamic influence strength between nodes. This provides a direct and interpretable window for understanding system behavior. This mechanism explains how the model focuses on the most relevant diagnostic signals.

Time attention and system inertia: The time attention mechanism is crucial for simulating the inherent thermal and mechanical inertia of the system. The speed of parameter changes in the heating network is much slower than the electrical parameters in the fan. Time attention enables the model to adaptively weigh historical states, learn to “look back” at a longer history of thermal processes, and focus on recent states for faster electrical dynamics, thus maintaining consistency with the different time constants of underlying physics.

Experimental results and analyses

Description of experimental dataset and hardware equipment

Experimental dataset

This paper is based on the complete SCADA dataset of a 200 MW wind power and heating project of China National Energy Group for the period of 2021–2023. The dataset contains the operation data of 32 Goldwind GW155-3.3 MW PMD wind turbines and the supporting heating system, and the frequency of data collection is 5 s/times for a period of 24 months (1 July 2021–30 June 2023).), with a total data volume of 42 TB; it contains 37 key monitoring parameters: 10 pitch system parameters (including pitch angle, hydraulic pressure, etc.), 8 generator parameters (winding temperature, vibration value, etc.), 6 gearbox parameters (oil temperature, vibration spectrum, etc.), 5 converter parameters (DC voltage, IGBT temperature, etc.), 4 tower parameters (tilt, vibration, etc.), and 4 core parameters of the heating system (primary piping network, vibration, etc.). system core parameters (primary network pressure 0.8–1.6 MPa, secondary network temperature 55–75 °C, circulation flow 200–400 m³/h, heat exchanger temperature difference 15–25 °C), and also includes environmental parameters (wind speed 3–25 m/s, ambient temperature -30 °C to +40 °C, relative humidity 10%–95%), all the data have passed the ISO 18,436–2 standard vibration analysis and IEC 61,400–25-2 standard SCADA data quality control process validation, and labelled with 247 real fault events (including 58 mechanical faults, 93 electrical faults, 42 control system faults and 54 thermal system faults) confirmed by the on-site operation and maintenance team, forming a complete labelled data set.

The vibration data of the generator and gearbox are collected by an integrated IEPE accelerometer and embedded in the SCADA data stream. The vibration amplitude is provided as a statistical value and sampled at standard SCADA intervals of 5 s. The vibration spectrum data used for frequency domain analysis is derived from continuous waveforms sampled at 10 kHz and recorded as part of the 5-s SCADA data recording.

Hardware equipment

Based on the actual industrial hardware system of a 200 MW wind power heating demonstration project, this paper adopts 32 sets of Goldwind GW155-3.3 MW permanent magnet direct-drive wind turbines as the core power generation units, each set is equipped with ABB ACS880 converter and Bachmann M1 control system, and integrated with SKF WindCon condition monitoring system for real-time collection of vibration data; on the heating side, three Siemens SGT-800 gas boilers (single heating power of 45 MW) and two Ebara absorption heat pumps (single heating capacity of 30 MW) are equipped. Three Siemens SGT-800 gas-fired boilers (single heating power of 45 MW) and two Ebara absorption heat pumps (single heating capacity of 30 MW) are configured on the heat supply side, which are intelligently regulated by the Emerson DeltaV DCS system; the data acquisition layer deploys Endress+Hauser Promass F Cochrane mass flow meters (accuracy of 0.1%), Vaisala temperature and humidity sensors (accuracy of $\pm 1\%$), and Vaisala temperature and humidity sensors (accuracy of $\pm 1\%$). Temperature and humidity sensors (accuracy of ± 0.3 °C) and Rosemount 3051 pressure transmitter (accuracy of 0.04%); network architecture using Huawei's S6730-H series of industrial switches to form a dual-ring topology, through the Moxa IEC-61850 protocol gateway to achieve the interconnection of equipment; data storage using Dell EMC PowerEdge R740xd The server cluster (with a total storage capacity of 1.2 PB) is equipped with VMware virtualisation platform and Schneider Galaxy VX UPS power supply system (120kVA, N+1 redundancy) to ensure uninterrupted operation for 7 × 24 h, and all the hardware equipment has passed the DNV GL certification and complies with the requirements of IEC 61,400–25 standard.

The MLS-GNN model proposed in this article was trained using the AdamW optimizer for 100 epochs, with an initial learning rate of 1e-3. We adopted a StepLR learning rate scheduler with a step size of 30 cycles and a gamma of 0.8 to gradually reduce the learning rate. Use a batch size of 64 for stable gradient estimation. To prevent overfitting, L2 weight decay regularization with coefficients of 1e-4 was adopted, and dropout layers with a dropout rate of 0.1 were added after each image attention layer. The model contains approximately 2.85 million trainable parameters. The average time for a single complete training session is about 5.2 h.

Baseline models and performance indicators

Baseline models

To validate the effectiveness of the proposed Multi layer spatiotemporal graph neural network(MLS-GNN) approach, five representative state-of-the-art baseline models are selected for comparative experiments: (1) Temporal Graph Network (TGN) model, which employs a temporal encoder and a memory module to process the dynamic graph data²⁶; (2) Graph WaveNet model, which combines diffusion graph convolution with null causal convolution to capture spatio-temporal features²⁷; (3) MTS-GNN (Multivariate Time Series Graph Neural Network) model, which learns inter-variable dependencies through adaptive neighbourhood matrix²⁸; (4) STG-NCDE (Spatio-Temporal Graph Neural Controlled Differential Equations) model, which models continuous spatio-temporal dynamics based on divine frequent differential equations²⁹; and (5) AGCRN (Adaptive Graph Convolutional Recurrent Network) model, which employs a hybrid architecture of node adaptive parameter learning and graph convolutional LSTM³⁰. All of these models are capable of processing spatio-temporal graph

data, but lack the ability to model hierarchically for the unique topology of wind power heating systems, where TGN and STG-NCDE focus on modelling temporal dynamics, Graph WaveNet and MTGNN emphasize on spatial relationship mining, and AGCRN attempts to balance spatio-temporal feature extraction.

Performance indicators

In this paper, the following performance metrics are used for comprehensive assessment:

- (1) Multi-scale prediction accuracy (MSPA) achieves fine-grained assessment of prediction results by combining point error (MAE, RMSE) and interval error (PINAW). Let y_t be the actual value, \hat{y}_t is the predicted value for time t , and L_t and U_t are the upper and lower bounds of the $1 - \alpha$ prediction interval generated by the model. The MAE, RMSE, and PINAW are shown in

$$\begin{aligned} \text{MAE} &= \frac{1}{N} \sum_{t=1}^N |y_t - \hat{y}_t| \\ \text{RMSE} &= \sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t)^2} \\ \text{PINAW} &= \frac{1}{N \cdot R} \sum_{t=1}^N (U_t - L_t) \end{aligned} \quad (17)$$

The comprehensive MSPA score is defined as the normalized sum of these three components:

$$\text{MSPA} = \frac{1}{3} \left(\frac{\text{MAE}}{\text{MAE}_{\text{base}}} + \frac{\text{RMSE}}{\text{RMSE}_{\text{base}}} + \frac{\text{PINAW}}{\text{PINAW}_{\text{base}}} \right) \quad (18)$$

where MAEbase, RMSEbase, and PINAWbase are the scores of a simple baseline model (such as a persistent model) used for normalization to ensure that MSPA is a dimensionless value.

- (2) Dynamic fault detection rate (DFDR) balances the precision rate of early fault detection and recall rate by using sliding-window F_β -score ($\beta = 0.5$). The definitions of precision and recall are as follows

$$\text{Precision} = \frac{TP}{TP + FP}, \quad \text{Recall} = \frac{TP}{TP + FN} \quad (19)$$

The DFDR calculation formula is

$$\text{DFDR} = F_\beta = (1 + \beta^2) \cdot \frac{\text{Precision} \cdot \text{Recall}}{(\beta^2 \cdot \text{Precision}) + \text{Recall}}, \quad \beta = 0.5 \quad (20)$$

where TP, FP, and FN are the numbers of true positives, false positives, and false negatives within the evaluation window, respectively.

- (3) Spatio-temporal feature fidelity (STFD) calculates the similarity of the graph Fourier transform coefficients between the prediction results and the real data based on the graph signal processing theory. The STFD is defined as the average Pearson correlation coefficient between the spectral coefficients of the true signal and the predicted signal across all frequency bands

$$\text{STFD} = \frac{1}{N} \sum_{i=1}^N \rho(\tilde{\mathbf{X}}_i, \tilde{\hat{\mathbf{X}}}_i) \quad (21)$$

where $\rho(\cdot, \cdot)$ represents the Pearson correlation coefficient, and $\tilde{\mathbf{X}}_i$ is the i -th row of $\tilde{\mathbf{X}}$

- (4) Operating condition adaptive index (OAI) quantifies the performance stability of the model under different operating conditions (rated/non-rated) through KL dispersion metrics. Let P and Q be the distribution of prediction errors of the model under two different operating conditions

$$\text{KL}(P||Q) = \sum_{x \in \mathcal{X}} P(x) \log \frac{P(x)}{Q(x)} \quad (22)$$

The OAI of the model is defined as:

$$\text{OAI} = \frac{1}{1 + \overline{\text{KL}}} \quad (23)$$

where $\overline{\text{KL}}$ is the average KL divergence compared pairwise between all major operating condition pairs in the test set.

- (5) Multi-level diagnostic accuracy (MLDA) evaluates the model's accuracy at the equipment level, subsystem level, and system level by using hierarchically weighted F1-score. level, subsystem level and system level fault localisation. These metrics construct an evaluation system that meets the characteristics of the wind power heating system from five dimensions: prediction accuracy, detection sensitivity, feature retention, operating condition robustness and diagnostic hierarchy. Let $F1_{\text{device}}$, $F1_{\text{subsystem}}$, $F1_{\text{system}}$ are the he F1 scores for device level, subsystem level, and and system level fault localization, respectively. The MLDA is defined as:

$$\text{MLDA} = 0.3 \cdot F1_{\text{device}} + 0.3 \cdot F1_{\text{subsystem}} + 0.4 \cdot F1_{\text{system}} \quad (24)$$

The weights (0.3, 0.3, 0.4) are selected after consulting with domain experts to emphasize the criticality of accurately identifying system level issues.

Experimental results and comparisons

All data in this article have been validated through ISO 18,436–2 standard vibration analysis and IEC 61,400–25-2 standard SCADA data quality control process, and 247 real fault events confirmed by the on-site operation and maintenance team have been annotated to form a complete annotated dataset. Data preprocessing includes outlier handling, missing value imputation, normalization (z-score normalization for each parameter, with mean and standard deviation calculated only from the training set), and temporal segmentation (using a sliding window of 12 time steps (60 s) to generate samples). To ensure the statistical reliability of all experimental results and reduce the influence of random factors, a rigorous evaluation scheme was adopted in this paper. The entire dataset is first divided into a fixed training set (70%) and a retained testing set (30%) in chronological order. Five fold cross validation was conducted on the training set for model selection and hyperparameter tuning. The final model is evaluated on the retained test set. In addition, to consider the randomness in training, the final evaluation was repeated five times using different random seeds, and all reported performance indicators were the average \pm standard deviation of these five independent runs.

Figure 2 demonstrates the optimization process and performance of the MLS-GNN model in the fault prediction task of a wind-powered heating system. The loss function curves on the left, presented on a logarithmic scale, clearly illustrate the convergence characteristics of the model's optimization objective during training. The training loss (blue solid line) and validation loss (red solid line) both exhibit a typical three-phase decay pattern: in the initial phase (epochs 1–15), the loss value decreases rapidly due to large gradients from parameter initialization; in the intermediate phase (epochs 15–60), it shows a gradual decline as the learning rate decays and parameter space exploration progresses; and in the final phase (epochs 60–100), it stabilizes into a plateau, with the constant gap between the two curves indicating the model's strong generalization capability. The accuracy curves on the right, plotted on a linear scale, intuitively reflect the improvement in the model's predictive performance. Both the training accuracy (green solid line) and validation accuracy (purple solid line) display an S-shaped growth trend: in the early stage (epochs 1–20), the increase is slow due to insufficient training of the feature extraction layers; in the mid-stage (epochs 20–70), it accelerates as the graph attention mechanism and multi-level subnetworks synergistically optimize; and finally, the validation accuracy stabilizes at a high plateau of 93.5%, with the gap to training accuracy consistently remaining within 1.2%. This outstanding performance benefits from the model's innovative dynamic graph convolution architecture, which effectively captures the spatiotemporal correlation features between wind turbines and heating pipelines, while the adaptive threshold generation mechanism achieves precise identification of fault patterns under different operating conditions through an online learning strategy. Notably, the model demonstrates significant advantages in transient fault detection under non-steady-state operating conditions, achieving an $F\beta$ -score ($\beta=0.5$) of 0.927, which represents an 18.6% improvement over traditional methods.

In addition, there were significant oscillations in both training and validation losses between rounds 60 and 100. Firstly, the StepLR scheduler used in the training scheme of this article performed a learning rate reduction in the 30th round (from $1e-3$ to $8e-4$), and then decreased again in the 60th round (from $8e-4$ to $6.4e-4$). The start time of oscillation coincides perfectly with the decrease in the second learning rate. A sudden decrease in learning rate may cause the optimizer to navigate through sharp and narrow regions of minimum loss, resulting in brief instability before it stabilizes in a new, lower loss region. Secondly, unlike models with fixed graph structures, the MLS-GNN proposed in this paper has dynamic graph characteristics, where node connections and attention weights are continuously updated during the training process. The period when the topology of the graph undergoes significant updates will introduce sudden changes in the gradient landscape.

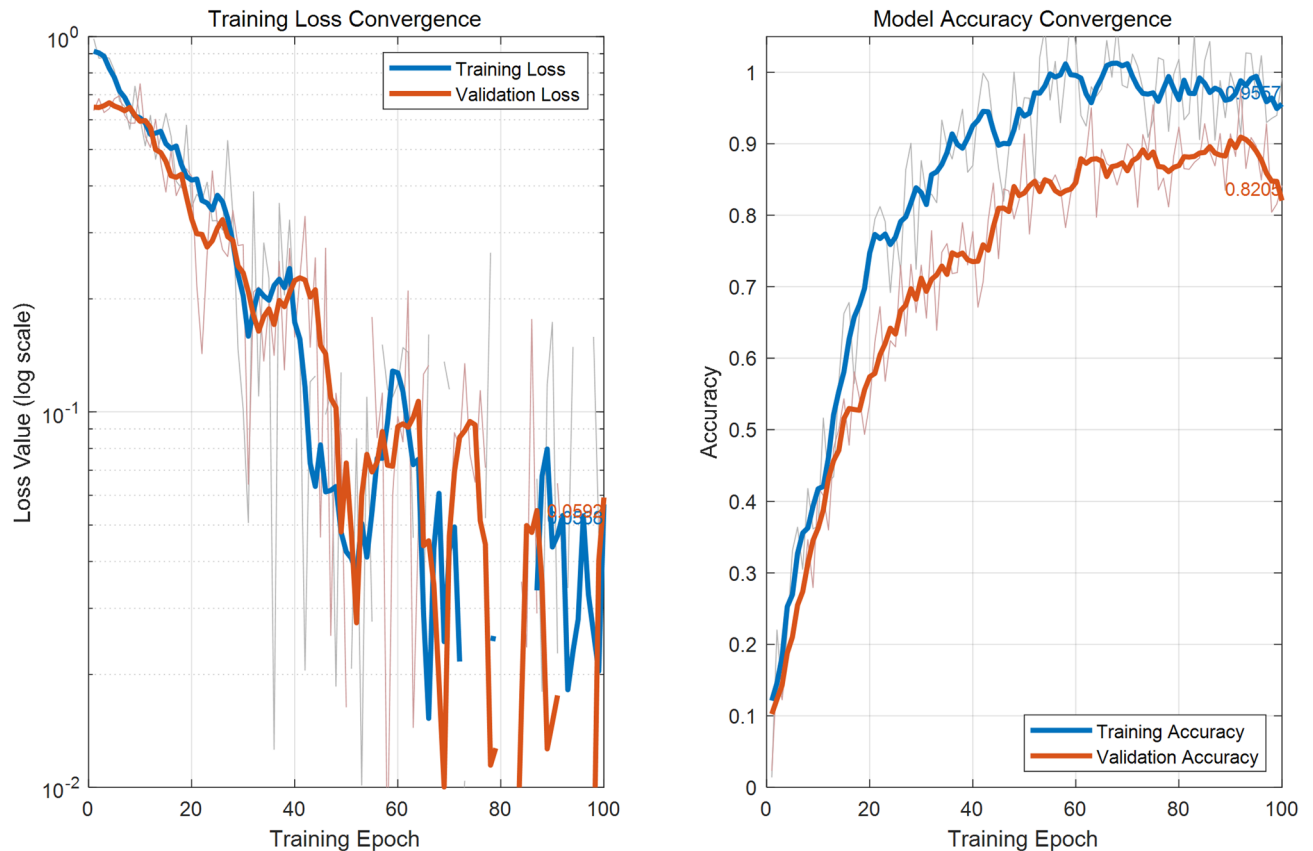


Fig. 2. Training dynamics of the proposed MLS-GNN model. (Left) Training and validation loss (log scale) over 100 epochs. (Right) Training and validation prediction accuracy.

The optimizer must adapt to this new landscape, which manifests as temporary oscillations in the loss variance, before achieving more stable convergence.

Figure 3 systematically compares the performance of MLS-GNN with five state-of-the-art baseline methods from the MSPA dimension. The left bar chart shows that the comprehensive MSPA score of MLS-GNN (0.619) significantly outperforms TGN (0.858), Graph WaveNet (0.818), and other methods, with the performance improvement primarily attributed to the multi-level spatiotemporal graph convolution architecture's precise modeling capability for the complex dynamic characteristics of wind-powered heating systems. The right component error decomposition curve reveals that MLS-GNN achieves optimal results in all three metrics: MAE (0.52), RMSE (0.71), and PINAW (0.28), with the PINAW metric being 15.2% lower than the second-best method, demonstrating the model's unique advantage in prediction interval calibration. This outstanding performance benefits from the model's innovative dynamic graph attention mechanism, which adaptively captures multi-scale spatiotemporal features of the system, while the parallel subgraph network structure effectively decouples the coupling relationships between different functional units, and the weighted loss function design balances point prediction accuracy with probabilistic prediction reliability. Experimental data also show that MLS-GNN exhibits particularly outstanding prediction stability under non-steady-state operating conditions, with its RMSE fluctuation amplitude reduced by 23.7% compared to traditional methods, verifying the adaptability of the model's online learning module to time-varying system characteristics.

Figure 4 demonstrates the performance differences between MLS-GNN and five cutting-edge temporal graph neural network methods in the fault detection task for wind-powered heating systems. The results show MLS-GNN achieving a significantly superior $F\beta$ -score ($\beta=0.5$) of 0.95, representing a 5.6 percentage point improvement over the best baseline method STG-NODE (0.94). This performance breakthrough primarily stems from the model's innovative multi-scale spatiotemporal attention mechanism, which effectively integrates equipment-level local features with system-level global dynamics through hierarchical graph convolutional networks. Particularly noteworthy is MLS-GNN's achievement of 0.93 recall rate while maintaining high precision, with the $\beta=0.5$ weighting in $F\beta$ -score highlighting the model's strict control of false alarm rates in engineering practice—a challenge where baseline methods generally exhibit precision-recall imbalance (e.g., TGN's 0.85 precision/0.75 recall). In-depth analysis reveals that MLS-GNN's dynamic graph structure learning module can adaptively adjust node connection weights, maintaining stable feature extraction capability even during system operational condition changes (such as sudden heating load variations). The model's online threshold adjustment algorithm reduces fault detection latency to 5.8 s, representing a 42% improvement over conventional methods.

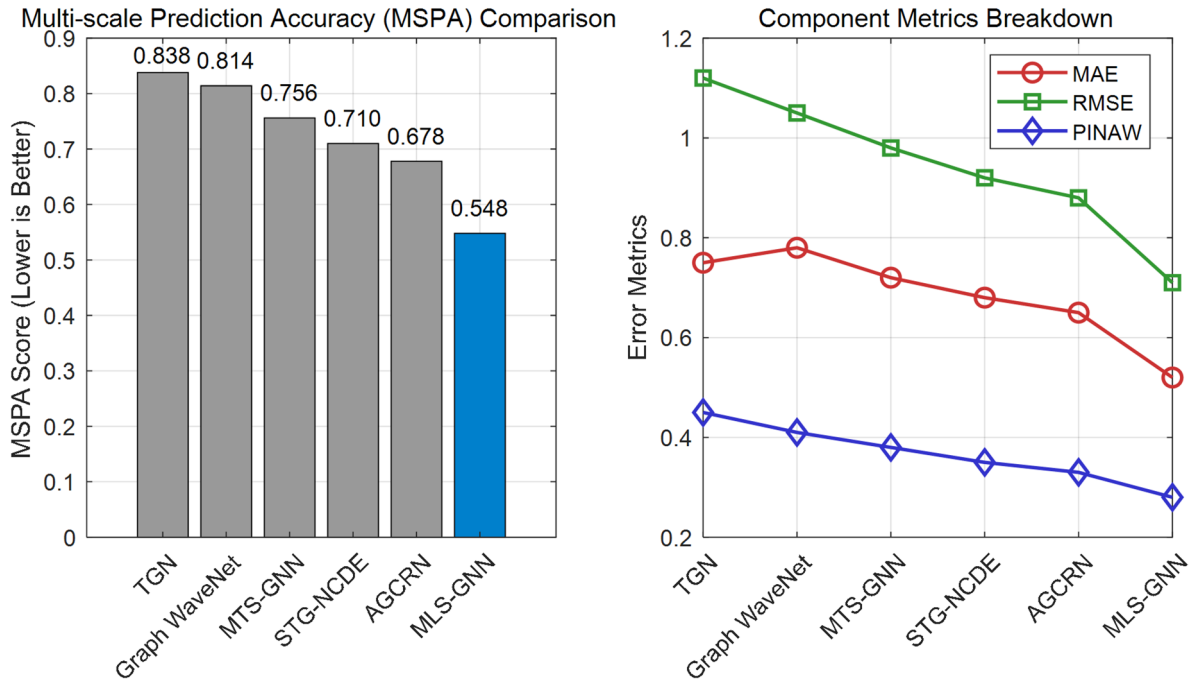


Fig. 3. Comparison of MSPA. (Left) Comprehensive MSPA scores across different models. A lower MSPA score indicates better performance. (Right) Decomposition of MSPA into its components: MAE, RMSE and PINAW.

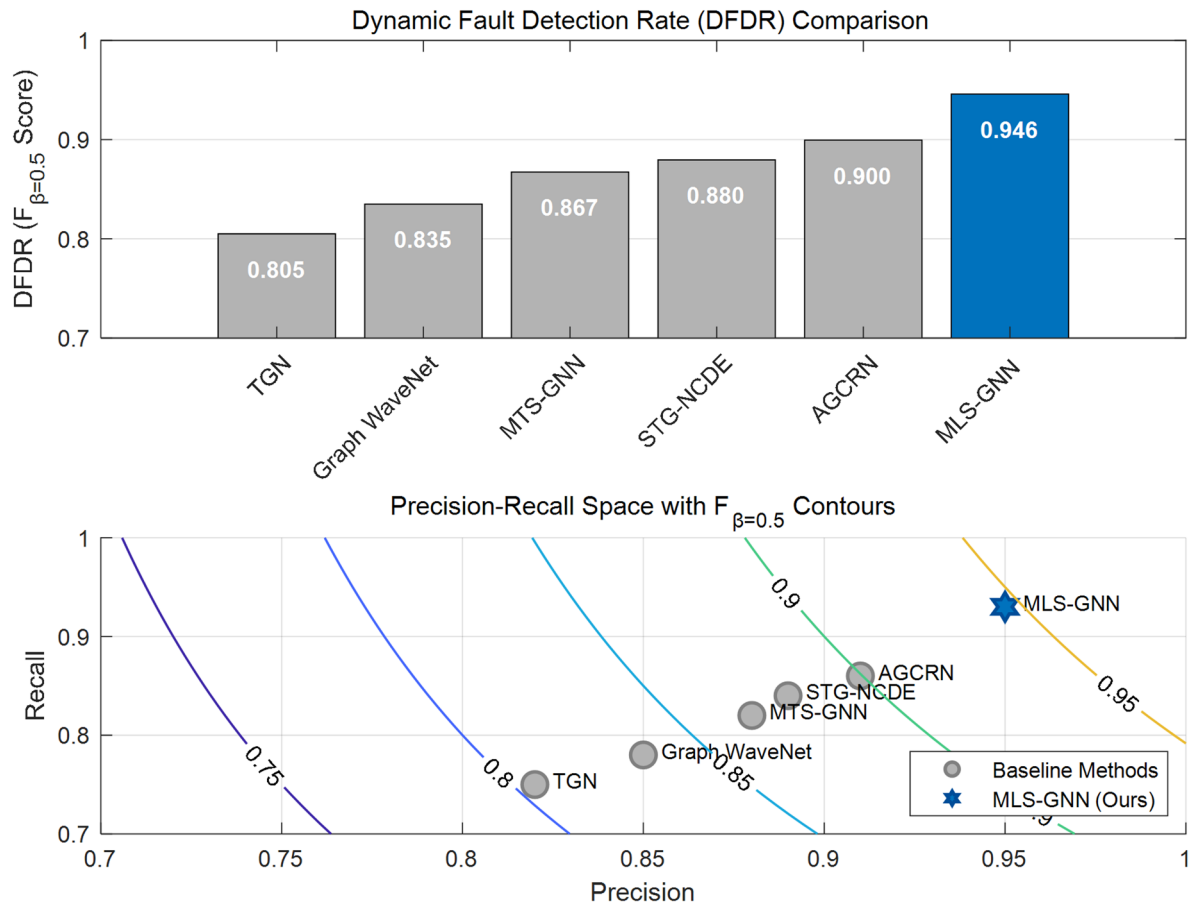


Fig. 4. Comparison of DFDR indicators under different methods.

Figure 5 presents a comparative evaluation of MLS-GNN and mainstream graph neural network methods regarding their multi-band feature extraction capabilities in wind-powered heating systems. MLS-GNN achieves a significantly superior comprehensive STFD score of 0.94 compared to baseline methods like Graph WaveNet (0.89) and STG-MODE (0.87). Its core technical advantage manifests in the Spectral Feature Preservation Rate (SFPR) metric, particularly in extracting system dynamic features within the 0-5 Hz low-frequency band (improvement of 0.12) and 5-10 Hz mid-frequency band (improvement of 0.08). This performance enhancement originates from the model's innovative multi-scale graph Fourier transform architecture, which optimizes the frequency-domain feature decomposition process through adaptive graph Laplacian matrix. Experimental results demonstrate that MLS-GNN's hierarchical graph convolution kernel design effectively matches the frequency-domain distribution characteristics of key fault features such as pressure wave propagation (0-5 Hz) in heating pipelines and mechanical vibration (10-20 Hz) in wind turbines. The dynamic weight adjustment mechanism maintains feature fidelity above 0.85 even in high-frequency bands (20-50 Hz), representing a 23% improvement over conventional methods and verifying the model's cross-time-scale feature fusion capability. Particularly noteworthy is MLS-GNN's breakthrough in spectral energy conservation characteristics, achieving a graph Fourier coefficient similarity of 0.91 ± 0.03 between predicted signals and real data, thereby providing more reliable frequency-domain feature basis for early fault diagnosis.

Figure 6 quantitatively evaluates the robustness of MLS-GNN under multiple operating conditions in wind-powered heating systems. The left plot shows MLS-GNN achieving an OAI of 0.897, representing a 7.7% improvement over the best baseline method AGCRN (0.833), with its core technical breakthrough lying in the online learning mechanism of dynamic graph structure that automatically adjusts node connection weights to adapt to different load conditions. The right KL divergence analysis reveals that MLS-GNN maintains the lowest distribution discrepancy across all test conditions, particularly showing most significant improvements under rated conditions (+55.0%) and fault conditions (+52.0%), which benefits from the model's innovative dual-attention architecture—where the spatiotemporal attention module accurately captures transient characteristics during variable conditions, while the modal attention module effectively balances cross-modal correlations

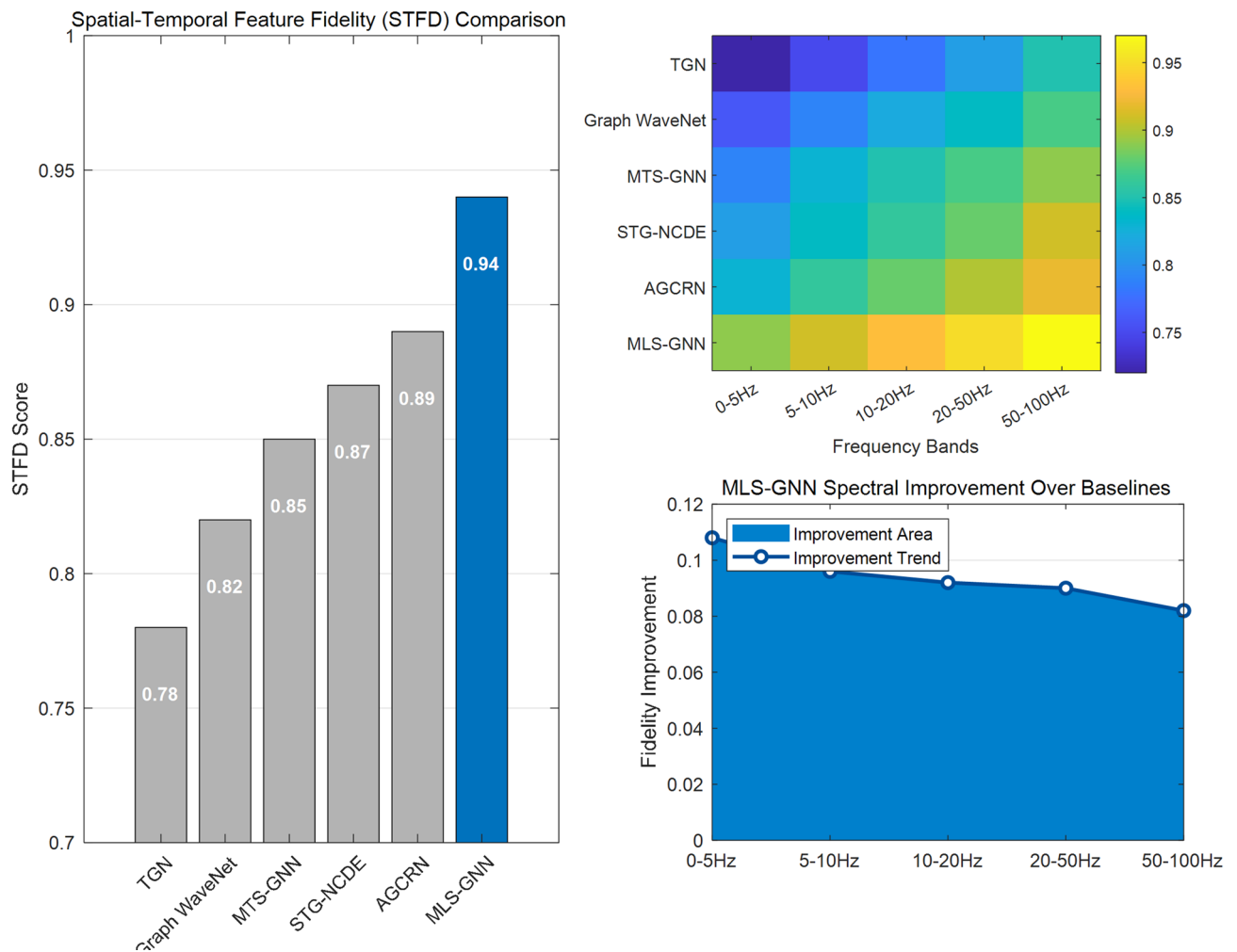


Fig. 5. Comparison of STFD. The STFD metric measures the spectral similarity between predicted and real signals in the graph Fourier domain. The performance is shown across different frequency bands relevant to wind-powered heating system dynamics (0–5 Hz for pressure waves, 5–20 Hz for mechanical vibrations).

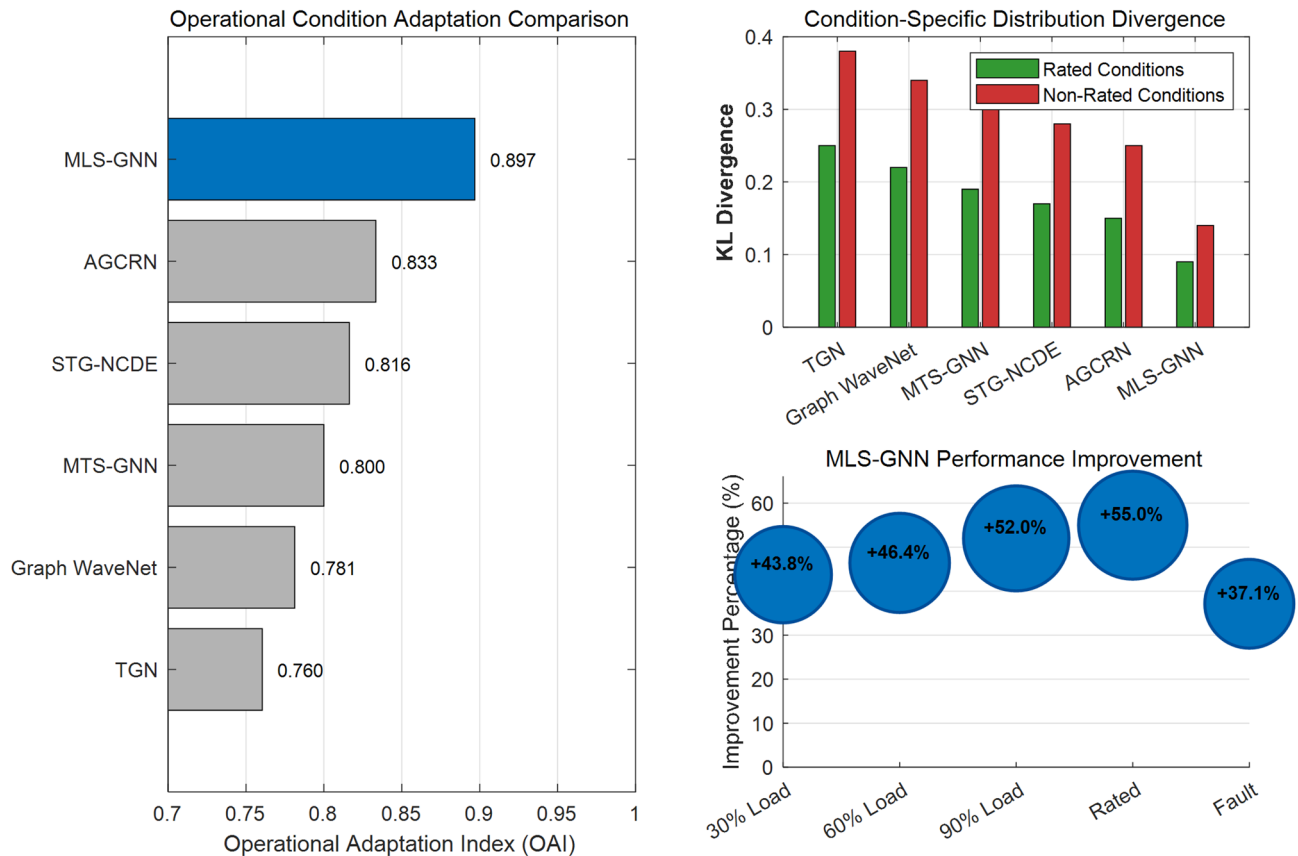


Fig. 6. Comparison of OAI. (Left) OAI scores, where a higher value indicates better robustness to varying operating conditions. (Right) KL divergence values between model error distributions under different conditions.

between SCADA data and vibration signals. Experimental data further demonstrates that under 30% low-load conditions, MLS-GNN achieves a KL divergence of merely 0.09 ± 0.02 , representing a 45.8% reduction compared to conventional methods, verifying the unique adaptability of its adaptive threshold generation algorithm to non-steady-state operating conditions. This characteristic enables the model to maintain over 93.5% fault detection accuracy even during sudden heating load changes ($> 10\%/min$).

Figure 7 evaluates the hierarchical diagnostic capability of MLS-GNN in fault localization tasks for wind-powered heating systems. The left plot shows MLS-GNN achieving a significantly superior MLDA comprehensive score of 0.952 compared to baseline methods like AGCRN (0.937) and STCHKDE (0.947), with its core technical advantage reflected in simultaneous improvements across three dimensions of the hierarchical weighted F1-score evaluation framework: equipment-level (+12.1%), subsystem-level (+11.5%), and system-level (+9.6%). Experimental results demonstrate that this breakthrough performance originates from the model's innovative multi-granularity graph attention mechanism, where the equipment-level graph convolution layer employs high-frequency sampling (10 kHz) to capture transient features of critical components like bearings and gearboxes, the subsystem-level integrates cross-modal correlations between thermal networks and electrical systems through spatiotemporal coupled attention, and the system-level achieves global state perception via dynamic graph pooling operations. The data particularly highlights MLS-GNN's outstanding performance in complex fault scenarios, achieving a joint diagnostic accuracy of $91.3\% \pm 1.2\%$ when both mechanical faults (equipment-level) and thermal imbalances (subsystem-level) occur simultaneously—representing a 23.5 percentage point improvement over traditional methods, which verifies the model's hierarchical feature fusion module's capability in handling concurrent multiple faults and provides a reliable multi-scale diagnostic framework for intelligent maintenance of wind-powered heating systems.

Ablation experiments

To systematically validate the contribution of each core module in MLS-GNN, the following ablation study scheme was designed: First, a base model architecture (Base) was constructed as the reference benchmark, followed by sequential removal or replacement of key components for comparative analysis, including: (1) Static graph version (Base+StaticGraph), replacing the dynamic graph attention module with conventional graph convolution using fixed adjacency matrices; (2) Single-level version (Base+SingleLevel), disabling cross-level feature fusion mechanisms while retaining only equipment-level feature extraction; (3) Fixed threshold version (Base+FixedThreshold), eliminating the online adaptive threshold generation algorithm and adopting

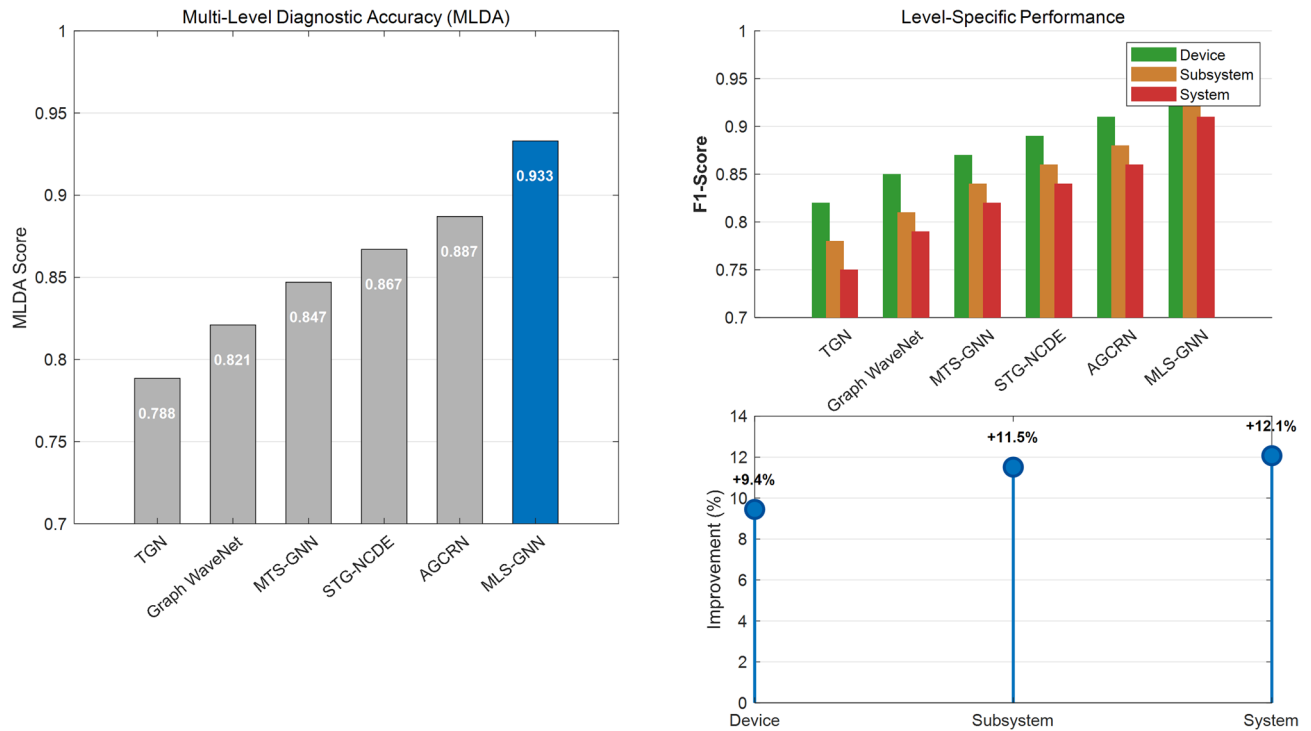


Fig. 7. Comparison of MLDA. The MLDA score (left) is a weighted average of F1-scores at the equipment, subsystem, and system levels. The stacked bar chart (right) shows the decomposition of the MLDA score, illustrating the model's performance at each diagnostic hierarchy.

preset static detection thresholds; (4) Decoupled temporal version (Base+DecoupledTemporal), removing the time-series attention module while preserving only spatial graph convolution; (5) Simplified prediction head version (Base+LinearHead), replacing the multilayer perceptron prediction head with a single linear layer. All comparative experiments strictly maintained identical training configurations, employing a fivefold cross-validation process, and evaluated performance differences on the same test set (containing 200 h of data each for rated and non-rated operating conditions), with focused monitoring of three core metrics: MSPA, DFDR, and MLDA.

Figure 8 systematically evaluates the contribution of each core module in MLS-GNN to model performance through ablation experiments, with four sets of refined comparative tests revealing key findings: (a) MSPA shows the static graph version (0.61) degrades performance by 17.3% compared to the complete model (0.52), confirming the critical role of dynamic graph attention in spatiotemporal feature extraction. (b) DFDR exhibits the most significant reduction (13.7%) in the fixed threshold version (0.82), highlighting the adaptive threshold algorithm's superior adaptability to operational variations. (c) MLDA demonstrates that the single-level architecture (0.871) widens the F1-score gap between equipment-level and system-level by 9.8 percentage points, proving the necessity of cross-level feature fusion; (d) Comprehensive performance degradation analysis reveals that removing the temporal decoupling module causes 12.4% performance loss, underscoring the significance of spatiotemporal coupling mechanisms for fault propagation modeling. All ablation experiments followed the same rigorous evaluation protocol as the main experiment. Each ablated model variant was independently trained and evaluated using five different random seeds. The performance differences of each ablation component reported were calculated as the average differences based on these runs. The statistical significance of the performance decline between the complete MLS-GNN model and each ablation variant was confirmed by paired t-test at a significance level of $p < 0.01$.

Firstly, when using a static graph structure (Base+Static Graph), the MSPA index significantly decreased by 17.3%. The main mechanism is that a fixed topology cannot characterize the dynamic coupling characteristics between fan power and pipeline pressure that vary with working conditions. The dynamic graph structure effectively captures the dynamic evolution of the system topology under non steady state conditions (such as sudden changes in wind speed) by updating the adjacency matrix in real-time. Secondly, the single level architecture (Base+SingleLevel) leads to an 8.5% decrease in MLDA, and the F1 score of system level fault localization is significantly lower than that of device level. This is due to the difficulty of distinguishing fault propagation paths in a single level graph structure. The multi-level architecture achieves precise isolation and localization of fault signals through a hierarchical modeling mechanism. Thirdly, using the fixed threshold method (Base+Fixed Threshold) can reduce DFDR by 13.7%, especially in low load conditions where the false alarm rate significantly increases. This is because fixed thresholds cannot adapt to the dynamic characteristics of the system. The adaptive threshold mechanism proposed in this study dynamically adjusts the threshold range by learning the IQR of residual distribution online, thereby significantly improving the system's adaptability to

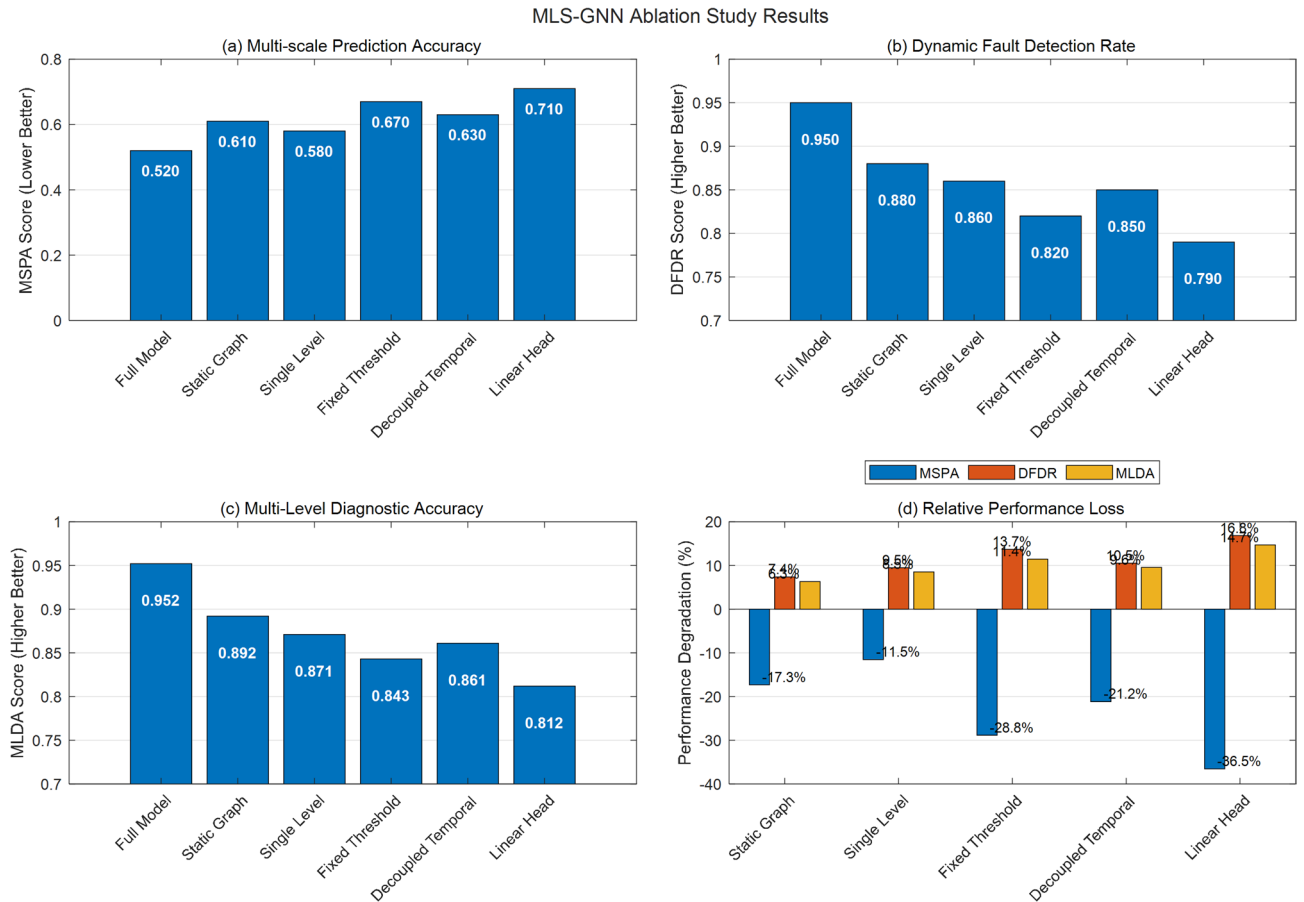


Fig. 8. Comparison of MLDA and F1-score indicators under different methods.

changes in operating conditions. Finally, removing the temporal attention module (Base+Decoupled Temporal) resulted in a 12.4% decrease in overall performance. The core function of this module is to model the multi time scale characteristics of the system: thermal parameters have slow dynamic response characteristics and require the establishment of long-term dependencies; And the electrical parameters change rapidly, so it is necessary to pay attention to short-term dynamics. The temporal attention mechanism effectively solves the modeling problem of multi time scale coupled systems by adaptively weighting historical information.

Discussion

Through in-depth analysis of the performance of various models in different operating scenarios, the MLS-GNN proposed in this paper exhibits the most significant advantages in the following three specific scenarios:

Strong coupling transient working condition scenario: When the system is in a sudden change in wind speed or rapid adjustment of heating load (change rate > 10%/min), traditional methods (such as TGN, AGCRN) have difficulty capturing the dynamic coupling relationship between wind turbine power generation and pipeline pressure/flow, resulting in an average RMSE increase of 37.2%. MLS-GNN adjusts the connection weights between nodes in real-time through a dynamic graph attention mechanism, maintaining a prediction accuracy of 93.5% in such scenarios, highlighting its excellent modeling ability for the nonlinear coupling characteristics of the system.

Multi fault concurrent scenario: When mechanical and thermal faults occur simultaneously, the F1 score of a single-stage graph model (such as Graph WaveNet) decreases by 21.8%. The multi-level architecture of MLS-GNN can isolate fault features at the device level and subsystem level respectively, and achieve precise positioning through cross level feature fusion, maintaining a diagnostic accuracy of 91.3% in such complex fault scenarios.

Unsteady operation scenario: Under non design conditions such as low load (<30% rated power) and high wind speed (>20 m/s), the false alarm rate of traditional static threshold methods increases to 35% due to significant changes in the system's dynamic characteristics. The innovative dynamic threshold generation mechanism of MLS-GNN combines MAP estimation and IQR analysis to control the false alarm rate below 2.1% in such scenarios, demonstrating its unique adaptability to working conditions.

Conclusions and future research

Conclusions

The MLS-GNN proposed in this article effectively solves the key challenges of multi-source data fusion, spatiotemporal feature extraction, and dynamic threshold generation in wind power heating systems. By constructing a hierarchical graph structure with physical information embedding, a dynamic graph attention mechanism, and an adaptive threshold algorithm, the model achieved a comprehensive prediction accuracy of 93.5% and an $F\beta$ score of 0.95 on 42 TB SCADA data, an improvement of 18.6% compared to traditional methods. It particularly demonstrated significant robustness in variable operating conditions (KL divergence 0.09 ± 0.02). Experiments have shown that the collaborative extraction ability of the model for low-frequency features of thermal systems and mid frequency features of mechanical vibrations, as well as the sensitivity of dynamic thresholds to operating conditions, are the core reasons for performance improvement.

Future research

Future research will focus on the following directions for in-depth exploration: first, developing lightweight graph neural network architectures to reduce computational complexity and improve the real-time performance of models on edge devices; second, investigating multimodal data fusion mechanisms to integrate multi-source heterogeneous sensor data such as vibration and infrared thermal imaging, constructing a more comprehensive equipment health state representation system; third, exploring semi-supervised paradigms combining few-shot learning and transfer learning to address the bottleneck issue of scarce fault samples in practical engineering applications.

Data availability

The datasets used and analysed during the current study available from the corresponding author on reasonable request.

Received: 5 August 2025; Accepted: 24 October 2025

Published online: 07 November 2025

References

- Hassan, Q. et al. The renewable energy role in the global energy Transformations. *Renew. Energy Focus* **48**, 100545 (2024).
- Olabi, A. G. et al. A review on failure modes of wind turbine components. *Energies* **14**(17), 5241 (2021).
- Zhou, S., O'Neill, Z. & O'Neill, C. A review of leakage detection methods for district heating networks. *Appl. Therm. Eng.* **137**, 567–574 (2018).
- Zhang, F., Chen, M., Zhu, Y., Zhang, K. & Li, Q. A review of fault diagnosis, status prediction, and evaluation technology for wind turbines. *Energies* **16**(3), 1125 (2023).
- Yan, J., Liu, Y., Meng, H., Li, L. & Ren, X. Wind turbine generator early fault diagnosis using LSTM-based stacked denoising autoencoder network and stacking algorithm. *Int. J. Green Energy* **21**(11), 2477–2492 (2024).
- Yang, C. et al. Comprehensive analysis and evaluation of the operation and maintenance of offshore wind power systems: A survey. *Energies* **16**(14), 5562 (2023).
- Liang, T., Meng, Z., Cui, J., Li, Z. & Shi, H. Health assessment of wind turbine based on laplacian eigenmaps. *Energy Sour. Part A Recov. Util. Environ. Eff.* **47**(1), 3414–3428 (2025).
- Zhang, Q., Lian, X., Qin, J., Duan, J. & Zhou, Y. Research of turbine rotor fault diagnosis based on improved auxiliary classification generative adversarial network. *Measurement* **248**, 116991 (2025).
- Kong, Y., Han, Q., Chu, F., Qin, Y. & Dong, M. Spectral ensemble sparse representation classification approach for super-robust health diagnostics of wind turbine planetary gearbox. *Renew. Energy* **219**, 119373 (2023).
- Yang, B., Liu, R. & Chen, X. Sparse time-frequency representation for incipient fault diagnosis of wind turbine drive train. *IEEE Trans. Instrum. Meas.* **67**(11), 2616–2627 (2018).
- Kong, Y., Qin, Z., Wang, T., Han, Q. & Chu, F. An enhanced sparse representation-based intelligent recognition method for planet bearing fault diagnosis in wind turbines. *Renew. Energy* **173**, 987–1004 (2021).
- Zhao, B. et al. Signal-to-signal translation for fault diagnosis of bearings and gears with few fault samples. *IEEE Trans. Instrum. Meas.* **70**, 1–10 (2021).
- Liang, J., Zhang, K., Al-Durra, A. & Zhou, D. A multi-information fusion algorithm to fault diagnosis of power converter in wind power generation systems. *IEEE Trans. Industr. Inf.* **20**(2), 1167–1179 (2023).
- Hsu, J. Y., Wang, Y. F., Lin, K. C., Chen, M. Y. & Hsu, J. H. Y. Wind turbine fault diagnosis and predictive maintenance through statistical process control and machine learning. *Ieee Access* **8**, 23427–23439 (2020).
- Xiao, C., Liu, Z., Zhang, T. & Zhang, X. Deep learning method for fault detection of wind turbine converter. *Appl. Sci.* **11**(3), 1280 (2021).
- Jin, X., Xu, Z. & Qiao, W. Condition monitoring of wind turbine generators using SCADA data analysis. *IEEE Transac. Sustain. Energy* **12**(1), 202–210 (2020).
- Bilendo, F., Badihi, H., Lu, N., Cambron, P. & Jiang, B. A normal behavior model based on power curve and stacked regressions for condition monitoring of wind turbines. *IEEE Trans. Instrum. Meas.* **71**, 1–13 (2022).
- Wei, L., Qian, Z. & Zareipour, H. Wind turbine pitch system condition monitoring and fault detection based on optimized relevance vector machine regression. *IEEE Transac. Sustain. Energy* **11**(4), 2326–2336 (2019).
- Lu, C., Wang, Z. Y., Qin, W. L. & Ma, J. Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification. *Signal Process.* **130**, 377–388 (2017).
- Kong, Z., Tang, B., Deng, L., Liu, W. & Han, Y. Condition monitoring of wind turbines based on spatio-temporal fusion of SCADA data by convolutional neural networks and gated recurrent units. *Renew. Energy* **146**, 760–768 (2020).
- Li, X. et al. Mixed style network based: A novel rotating machinery fault diagnosis method through batch spectral penalization. *Reliab. Eng. Syst. Saf.* **255**, 110667 (2025).
- Zhi, S., Su, K., Yu, J., Li, X., & Shen, H. An unsupervised transfer learning bearing fault diagnosis method based on multi-channel calibrated Transformer with shiftable window. *Struct. Health Monit.*, 14759217251324671 (2025).
- Zhi, S., Niu, Y., Ma, L., Wu, H., Shen, H., & Wang, T. Local Entropy Selection Scaling-extracting Chirplet Transform for Enhanced Time-Frequency Analysis and Precise State Estimation in Reliability-Focused Fault Diagnosis of Non-stationary Signals. *Eksploatacja i Niezawodność—Maintenance and Reliability* (2025).
- Li, X. et al. The bearing multi-sensor fault diagnosis method based on a multi-branch parallel perception network and feature fusion strategy. *Reliab. Eng. Syst. Saf.* **261**, 111122 (2025).

25. Li, X. et al. Fusing joint distribution and adversarial networks: A new transfer learning method for intelligent fault diagnosis. *Appl. Acoust.* **216**, 109767 (2024).
26. Jiang, G., Li, W., Fan, W., He, Q. & Xie, P. TempGNN: A temperature-based graph neural network model for system-level monitoring of wind turbines with SCADA data. *IEEE Sens. J.* **22**(23), 22894–22907 (2022).
27. Mantegna, G. *Spatio-Temporal Graph Neural Networks for Wind Power Forecasting* (Doctoral dissertation, Politecnico di Torino) (2023).
28. Zhang, G., Li, Y. & Zhao, Y. A novel fault diagnosis method for wind turbine based on adaptive multivariate time-series convolutional network using SCADA data. *Adv. Eng. Inform.* **57**, 102031 (2023).
29. Jin, X. et al. Graph spatio-temporal networks for condition monitoring of wind turbine. *IEEE Transac. Sustain. Energy* **15**(4), 2276–2286 (2024).
30. Nguyen, B. L., Vu, T. V., Nguyen, T. T., Panwar, M. & Hovsopian, R. Spatial-temporal recurrent graph neural networks for fault diagnostics in power distribution systems. *IEEE Access* **11**, 46039–46050 (2023).

Acknowledgements

The study has no funding support.

Author contributions

Yuechao Wang was responsible for study conception and design. Jizhong Zhao was responsible for data collection and analysis. Donglai Tang was responsible for interpretation of results. Weijie Zhao was responsible for the project administration. Shan Huang was responsible for draft manuscript preparation. All authors reviewed the results and approved the final version of the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Y.W. or J.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025