



Wind turbine drivetrain fault diagnosis and intelligent O&M: a review

Xueyi Li^{1,2}, Zige Wang¹, Wenyang Hu², Yun Kong³, Tianyang Wang², Zhaoye Qin², Fulei Chu²

Keywords:

Wind turbine drivetrain, fault diagnosis, deep learning, intelligent operation and maintenance, domain adaptation

Citation: Li, X.; Wang, Z.; Hu, W.; Kong, Y.; Wang, T.; Qin, Z.; Chu, F. Wind turbine drivetrain fault diagnosis and intelligent O&M: a review. *Intell. Robot.* 2026, 6(3), 750-95. <https://dx.doi.org/10.20517/ir.2026.33>

Received: 22 Jun 2026

First Decision: 22 Jul 2026

Revised: 4 Aug 2026

Accepted: 16 Sep 2026

Published: 30 Sep 2026

Academic Editor:

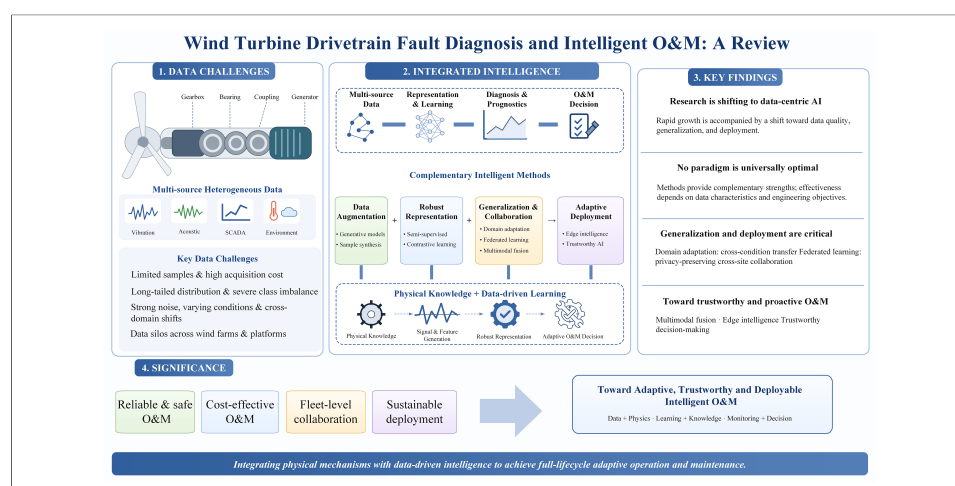
Chen Peng

Copy Editor:

Pei-Yun Wang

Production Editor:

Pei-Yun Wang



Abstract

As wind turbines evolve toward larger capacities, fleet-level clustering, and operation under complex conditions, fault mechanisms in key drivetrain components show multi-physics coupling and complex evolution, creating a major bottleneck in condition monitoring: models are often constructible but hard to generalize. Although deep learning is effective for end-to-end feature extraction, industrial data challenges - such as limited samples, long-tailed distributions, and domain shifts - severely restrict model generalization and engineering applicability. To address these issues, this paper systematically reviews fault diagnosis and intelligent operation and maintenance (O&M) technologies for wind turbine drivetrains. It analyzes generative-model-based sample augmentation methods and their limitations in authenticity and generalization improvement; examines the roles of semi-supervised, contrastive, and cost-sensitive learning in handling imbalanced and unlabeled data; and discusses how multimodal fusion, federated learning, and domain adaptation help alleviate data silos and environmental drift. Finally, it summarizes major challenges, including cross-domain generalization, explainability, and edge deployment, and proposes a future adaptive O&M paradigm integrating physical mechanisms with data-driven methods across the full asset lifecycle.

¹College of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin 150040, Heilongjiang, China.

²Department of Mechanical Engineering, Tsinghua University, Beijing 100084, China.

³School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China.

Correspondence to: Assoc. Prof. Yun Kong, School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China. E-mail: kongyun@bit.edu.cn; Assoc. Prof. Tianyang Wang, Department of Mechanical Engineering, Tsinghua University, Beijing 100084, China. E-mail: wty19850925@tsinghua.edu.cn

1. INTRODUCTION

Against the backdrop of the accelerated transition of the global energy structure toward decarbonization, wind energy has emerged as a crucial pillar of energy strategies due to its scalability and economic viability. As high-quality onshore resources near saturation, the wind power industry is rapidly shifting toward offshore development, turbine upscaling, and fleet clustering. However, wind turbines predominantly operate over extended periods under complex conditions characterized by strong randomness, intense physical coupling, and multiple disturbances. Consequently, critical drivetrain components - such as gearboxes, bearings, and couplings - are highly susceptible to degradation and failure. This not only severely compromises operational safety but also significantly elevates the full-lifecycle operation and maintenance (O&M) costs and the levelized cost of energy (LCOE). Therefore, establishing a highly reliable and strongly generalizable fault diagnosis and intelligent O&M system has become an urgent and critical bottleneck in wind power intelligent O&M.

In recent years, data-driven methods such as deep learning have demonstrated significant advantages in wind turbine fault diagnosis. They are capable of adaptively extracting nonlinear features from multi-source heterogeneous data, including vibration signals, acoustic signals, and supervisory control and data acquisition (SCADA) systems. SCADA data typically contain diverse operational and environmental variables, such as wind speed, active power, rotor speed, generator speed, gearbox temperature, bearing temperature, generator temperature, pitch angle, yaw position, and various electrical parameters, which provide essential information for characterizing turbine operating states and degradation processes^[1,2]. However, real-world engineering practices are plagued by several critical challenges: the high cost of fault sample acquisition, sample scarcity resulting from the low frequency of fault events, extreme class imbalance, and the inconsistent quality of multi-source data. This is starkly at odds with the data-hungry nature of deep learning algorithms. Consequently, models become highly susceptible to overfitting or bias toward majority classes, which significantly degrades generalization and deployment stability under intense noise, varying operating states, and distribution shifts.

To address the aforementioned challenges, existing research has primarily unfolded across three dimensions: data augmentation, learning paradigm innovation, and feature calibration. Specifically, generative models are employed to synthesize fault samples to bridge data gaps^[3,4]. However, they struggle to fully replicate authentic physical dynamics and risk training instability^[5]. Semi-supervised and contrastive learning enhance representation capabilities by mining the structural information of unlabeled data^[6,7], yet they remain highly sensitive to noise and distribution shifts^[8]. Cost-sensitive learning and resampling strategies can alleviate classification bias, but they yield marginal benefits under complex operating conditions characterized by highly overlapping features^[9]. Meanwhile, signal processing and feature enhancement form the foundation for improving robustness in few-shot scenarios^[10,11].

Furthermore, as the digital transformation of wind farms deepens, the data silo effect resulting from an exclusive reliance on single-turbine or single-station data has become increasingly prominent. Constrained by commercial privacy protections and data security barriers, high-value fault samples across different wind farms and operators are difficult to aggregate physically. Consequently, the industry is widely trapped in a paradox of being data-rich but knowledge-poor. Meanwhile, conventional pure data-driven models often lack physical consistency constraints. When confronted with unseen fault patterns or extreme weather disturbances, they are highly prone to generating predictions that violate physical and dynamic mechanisms. This black-box decision-making paradigm results in poor model explainability and high false-alarm rates, severely hindering trust and the large-scale deployment of artificial intelligence (AI) algorithms in safety-critical infrastructures like wind turbines. Therefore, breaking down data barriers, deeply integrating physical mechanism knowledge, and constructing a trustworthy fault diagnosis and intelligent O&M system equipped

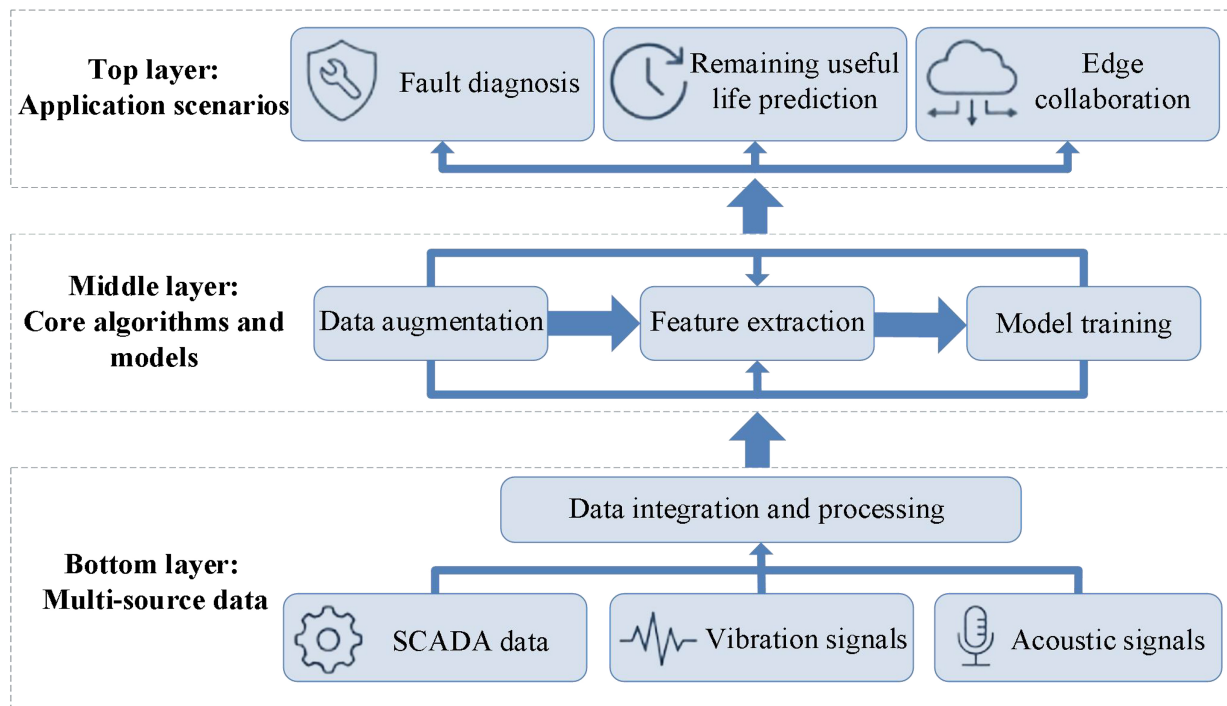


Figure 1. Framework for wind drivetrain fault diagnosis and smart O&M. O&M: Operation and maintenance; SCADA: supervisory control and data acquisition.

with privacy protection and continuous evolution capabilities have become an inevitable path for the evolution of next-generation intelligent diagnostic technologies. Oriented toward the full-lifecycle closed loop of fault diagnosis and intelligent O&M for key components of wind turbine drivetrains, this paper systematically reviews the critical technological spectrum spanning data acquisition and governance, algorithmic learning paradigms, multi-source collaboration, and engineering deployment.

To this end, this paper establishes the review framework illustrated in [Figure 1](#), systematically discussing the collaborative innovation mechanisms of wind power intelligent O&M technologies across the data, model, and application layers. From a data-augmentation perspective, it evaluates the efficacy of generative adversarial networks (GANs) and diffusion models for sample synthesis. Subsequently, at the algorithmic paradigm level, it provides an in-depth analysis of the mechanisms behind semi-supervised learning, contrastive learning, and transfer learning, with a particular focus on incremental learning and meta-learning strategies aimed at overcoming catastrophic forgetting. From the dimension of system collaboration, it explores federated learning (FL) and lightweight edge deployment pathways as solutions to the data silo problem. Finally, this paper envisions the construction of a next-generation intelligent O&M system that is jointly driven by data and physical mechanisms, and equipped with adaptive evolutionary capabilities.

2. STATISTICAL ANALYSIS OF LITERATURE

To gain a comprehensive understanding of the current research landscape in fault diagnosis and intelligent O&M for wind turbine drivetrain components - and particularly to explore the evolutionary trajectory of data-driven technologies in addressing the data dilemma - relying solely on qualitative summaries is often insufficient to capture the full scope of the literature. Therefore, this section conducts a quantitative scan and visual reconstruction of the relevant core literature published between January 2010 and June 2025. By establishing a systematic literature retrieval strategy, it demonstrates, across three dimensions-annual publication trends, keyword co-occurrence clustering, and technological evolution trajectories-how the

academic community has gradually transitioned from traditional physics-based models to intelligent technological systems represented by deep learning, transfer learning, and FL.

2.1. Literature search strategy and data sources

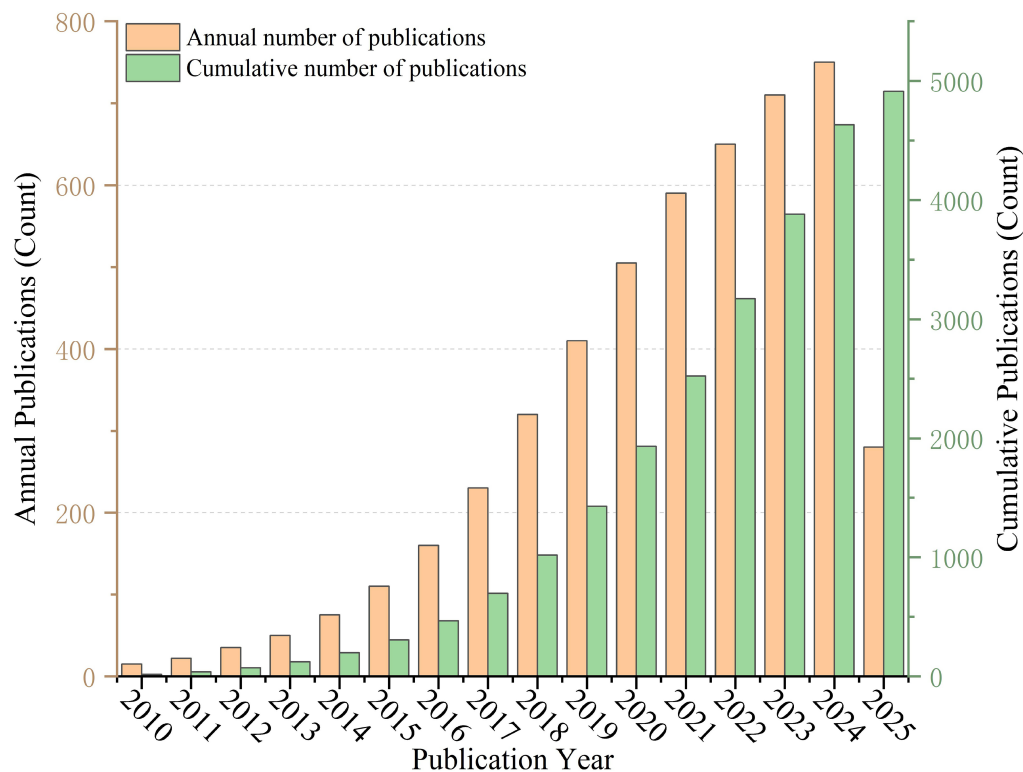
A structured literature search was conducted in the Web of Science Core Collection (<https://www.webofscience.com/wos/woscc/basicsearch>), IEEE Xplore (<https://ieeexplore.ieee.org/>), and Scopus (<https://www.scopus.com/>) to identify studies related to wind turbine fault diagnosis and intelligent O&M. All database searches were conducted in December 2025, while the publication period included in the bibliometric dataset was restricted to January 2010 - June 2025. Only English-language research articles and review articles were retained for bibliometric analysis. Publications outside the predefined time window, non-English publications, and document types other than research articles and review articles were excluded from the bibliometric dataset. This language restriction applies to the bibliometric dataset, not to all references cited in this review. A small number of relevant Chinese-language publications are cited as supplementary evidence for specific technical discussions and table entries. These publications are excluded from the 4,912-record bibliometric dataset and from the publication counts and keyword analysis reported in Section 2.

The search strategy was organized into three concept groups: research object, problem, and methodology. The object-related terms were “wind turbine”, “offshore wind farm”, and “drivetrain”; the problem-related terms were “fault diagnosis”, “imbalanced data”, “long-tail distribution”, “remaining useful life”, and “RUL”; and the methodology-related terms were “deep learning”, “transfer learning”, “generative adversarial network”, “GAN”, and “federated learning”. Terms within each concept group were combined using the Boolean operator OR, whereas the three concept groups were combined using AND. Searches were performed in the title, abstract, and keyword fields of the three databases.

For the Web of Science Core Collection, the search was implemented using the corresponding field codes for Title (TI), Abstract (AB), and Author Keywords (AK): (TI=("wind turbine*" OR "offshore wind farm*" OR drivetrain*) OR AB=("wind turbine*" OR "offshore wind farm*" OR drivetrain*) OR AK=("wind turbine*" OR "offshore wind farm*" OR drivetrain*)) AND (TI=("fault diagnosis" OR "imbalanced data" OR "long-tail distribution" OR "remaining useful life" OR RUL) OR AB=("fault diagnosis" OR "imbalanced data" OR "long-tail distribution" OR "remaining useful life" OR RUL) OR AK=("fault diagnosis" OR "imbalanced data" OR "long-tail distribution" OR "remaining useful life" OR RUL)) AND (TI=("deep learning" OR "transfer learning" OR "generative adversarial network*" OR GAN OR "federated learning") OR AB=("deep learning" OR "transfer learning" OR "generative adversarial network*" OR GAN OR "federated learning") OR AK=("deep learning" OR "transfer learning" OR "generative adversarial network*" OR GAN OR "federated learning"))).

For Scopus, the corresponding search was performed using the TITLE-ABS-KEY field: TITLE-ABS-KEY("wind turbine*" OR "offshore wind farm*" OR drivetrain*) AND TITLE-ABS-KEY("fault diagnosis" OR "imbalanced data" OR "long-tail distribution" OR "remaining useful life" OR RUL) AND TITLE-ABS-KEY("deep learning" OR "transfer learning" OR "generative adversarial network*" OR GAN OR "federated learning").

For IEEE Xplore, the same Boolean structure was applied through the Advanced Search interface using the Document Title, Abstract, and Author Keywords fields. The three conceptual groups were combined consistently as (Object terms) AND (Problem terms) AND (Method terms) to maintain comparable retrieval logic across the databases.



Note: 2025 annual publication volume is counted up to June.

Figure 2. Annual and cumulative counts of deduplicated literature records.

The initial searches returned 3,276 records from Web of Science Core Collection, 1,483 records from IEEE Xplore, and 3,047 records from Scopus, yielding a total of 7,806 records before deduplication. Publications released after June 2025, including publications appearing in 2026, were excluded from the bibliometric dataset.

Crossdatabase deduplication was implemented entirely within Zotero software following a standardized workflow:

All retrieved records from the three databases were imported into a single Zotero library for cross-database deduplication. Potential duplicate records were first identified using Zotero's Duplicate Items function and subsequently subjected to manual verification. DOI was used as the primary identifier when available. For records without a DOI, the article title, publication year, and author information were jointly compared to distinguish genuine duplicates from distinct but similar publications. Confirmed duplicate records were resolved in Zotero so that only one bibliographic record was retained for each duplicate set. After deduplication, the initial 7,806 records were reduced to 4,912 unique records, which constituted the final bibliometric dataset.

2.2. Annual publication trends and division into research stages

As illustrated in [Figure 2](#), the annual and cumulative counts of the 4,912 unique records obtained after cross-database deduplication in Zotero reveal a marked increase in both annual and cumulative publications concerning wind power intelligent O&M.

Based on the characteristics of technological evolution, the academic development trajectory of this field can be clearly divided into the following four stages:

1. The incubation and method transfer period (2010-2014): During this stage, the annual publication volume hovered at a low level with slow cumulative growth, indicating that the field was still in its exploratory phase. Research primarily focused on mechanism analysis and signal processing, complemented by shallow models such as linear discriminant analysis and k-nearest neighbors to verify feasibility.

2. The mechanism-driven and statistical learning-dominated period (2015-2017): Publications experienced a period of steady growth. The research paradigm predominantly leaned toward “feature engineering + shallow models”. The research focus concentrated on physical feature extraction based on signal processing, alongside traditional machine learning algorithms like support vector machines (SVM) and random forests. Although some studies began to recognize the prevalent data imbalance issue at industrial sites, proposed solutions were mostly confined to basic oversampling strategies such as Synthetic Minority Over-sampling Technique (SMOTE)^[12], and researchers had not yet deeply explored distribution patterns at the feature manifold level.

3. The deep learning explosion period (2018-2021): Accompanied by the maturation of deep neural network architectures such as convolutional neural networks (CNNs) and long short-term memory (LSTM) networks^[13], the publication volume in this phase exhibited an exponential surge. The core shift during this period was leveraging the end-to-end learning capabilities of deep models to gradually replace manual feature engineering that relied on expert experience. However, much research at this stage validated algorithm performance in ideal environments, paying insufficient attention to label-deficiency issues and distribution discrepancies between training and testing data in real-world wind farm settings.

4. The data-centric AI transition period (2022-June 2025): A significant problem-oriented shift occurred in research directions, transitioning from the mere pursuit of model accuracy to addressing the core pain points of engineering implementation. Recent studies have heavily focused on resolving challenges such as few-shot, zero-shot, and cross-domain generalization in industrial scenarios. Particularly in the latest literature from 2024 to 2025, the proportion of applications involving generative AI and FL has increased substantially. This marks a definitive shift in the industry’s research focus from pure algorithmic model optimization to the more profound challenges of overcoming data defects and enhancing model robustness.

2.3. Research hotspots and evolutionary trends in fault diagnosis technology

The keyword co-occurrence network was constructed in [Figure 3](#) using VOSviewer version 1.6.20 to examine the thematic relationships among terms related to intelligent O&M of wind turbines. Keywords were screened for relevance to the scope of this review, with particular emphasis on intelligent O&M, fault diagnosis, wind turbines, condition monitoring, predictive maintenance, and other closely related diagnostic topics. Rather than retaining keywords solely on the basis of a fixed occurrence-frequency threshold, thematic relevance to wind turbine fault diagnosis and intelligent O&M was adopted as the primary screening criterion. Therefore, no additional minimum-occurrence threshold was imposed after the thematic keyword-screening step.

The analysis type was set to co-occurrence, with all keywords as the unit of analysis. Full counting was adopted as the counting method. Network normalization was performed using the association-strength method. The clustering resolution was set to 1.00 and the minimum cluster size to 1. For network layout, the attraction and repulsion parameters were set to 2 and 0, respectively, while the remaining advanced layout and clustering parameters were retained at their default VOSviewer settings. No manual reassignment of

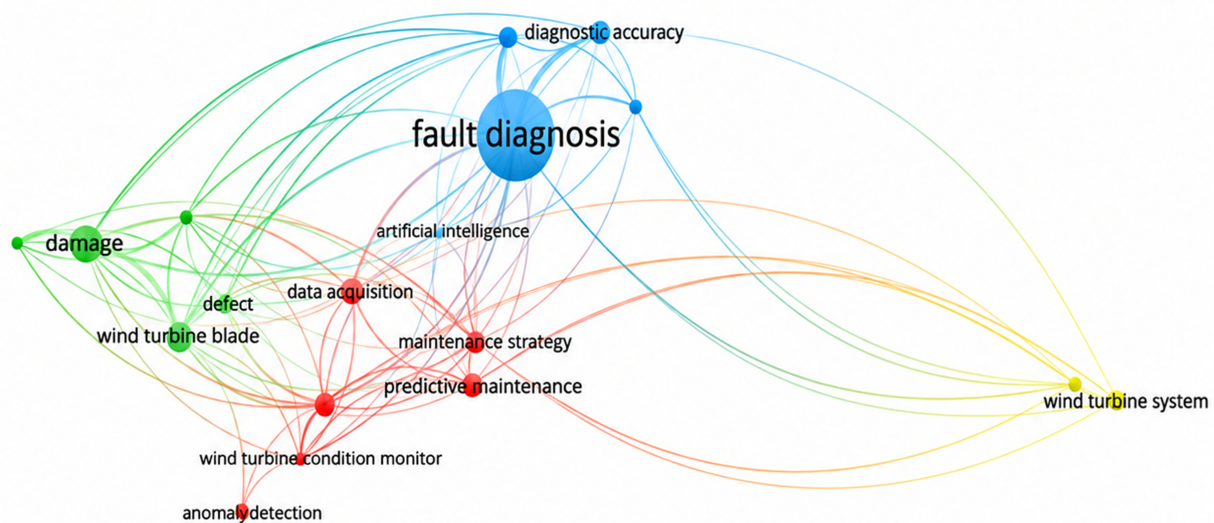


Figure 3. Keyword co-occurrence network map of wind turbine fault diagnosis.

keywords to clusters was performed. The resulting network contained 12 keyword nodes, 34 co-occurrence links, and four clusters. In the visualization, node size reflects keyword occurrence, link strength represents the strength of co-occurrence relationships between keywords, and node color indicates cluster membership.

On the left side of the network, the closely associated keywords “wind turbine blade”, “damage”, and “defect” form a thematic cluster related to component damage and defect identification, indicating sustained research attention to blade health monitoring and defect diagnosis.

In summary, fault diagnosis and intelligent O&M technologies for wind turbines are transitioning from reactive post-fault analysis to proactive predictive maintenance. Furthermore, digital twins, multimodal fusion, and edge intelligence are poised to become the dominant research directions in the coming years.

To facilitate reproducible and fair comparisons, the experimental datasets used in existing studies can be broadly divided into general rotating-machinery benchmarks, such as CWRU, Paderborn, XJTU-SY, and IMS, and wind-turbine-specific datasets, such as NREL gearbox test data and operational SCADA datasets^[14–18]. The former support controlled verification of fault classification, domain adaptation, and remaining useful life (RUL) prediction algorithms, whereas the latter better reflect environmental disturbances, sensor heterogeneity, and long-term operating variations in real wind farms. Model evaluation should also be matched to the specific task: classification studies should report precision, recall, macro-F1, or balanced accuracy in addition to overall accuracy; anomaly detection should consider false-alarm rate, missed-detection rate, area under the precision–recall curve (AUPRC), and detection delay; RUL prediction should use mean absolute error (MAE), root mean square error (RMSE), and prediction-horizon-related indicators; and lightweight deployment should jointly report parameter count, floating-point operations (FLOPs), model size, and inference latency. Because reported diagnostic-model performance depends strongly on dataset characteristics, class distributions, operating conditions, data-partitioning strategies, and hardware platforms, results from different studies should not be interpreted as direct performance rankings. A meaningful comparison should therefore consider not only the numerical values of evaluation metrics, but also the task definition, dataset representativeness, experimental protocol, and deployment constraints. [Table 1](#) summarizes the commonly used datasets, task-specific evaluation metrics, and minimum reporting requirements for more reproducible and transparent comparisons.

Table 1. Common experimental datasets and evaluation metrics in intelligent wind turbine O&M

Evaluation task	Dataset or metric focus	Recommended metrics	Main reporting requirement
Fault classification	CWRU bearing dataset	Accuracy, precision, recall, macro-F1	Report load condition, fault size, and known dataset anomalies ^[19]
Cross-condition diagnosis	Paderborn bearing dataset	Target-domain accuracy, average transfer accuracy	Clearly define source and target operating conditions ^[20]
RUL prediction	XJTU-SY bearing dataset	MAE, RMSE, MAPE, and R ²	Report training/test division and failure threshold ^[21]
Degradation monitoring	IMS bearing dataset	Health-index trend, MAE, RMSE, prediction horizon	Report the selected degradation starting point ^[22]
Wind turbine gearbox diagnosis	NREL GRC dataset	Accuracy and clustering-based diagnostic performance	Report selected sensor channels, operating conditions, and the labeling or clustering protocol ^[23]
Imbalanced anomaly detection	Precision–recall evaluation	AUPRC, precision, and recall	Prefer AUPRC when abnormal samples are extremely rare ^[24]
Imbalanced classification	MCC	MCC and balanced accuracy	Avoid relying only on accuracy or F1-score ^[25]
Generative augmentation (image evaluation)	Synthetic-image quality evaluation	FID, generative precision/recall, and classifier two-sample ROC-AUC	Specify image preprocessing and feature extraction; signal-domain validity requires separate assessment ^[26]
Edge deployment	Model computational efficiency	Parameters, FLOPs, model size, inference latency, and throughput	Specify hardware, input length, batch size, and precision ^[27]

O&M: Operation and maintenance; RUL: remaining useful life; MAE: mean absolute error; RMSE: root mean square error; MAPE: mean absolute percentage error; R²: coefficient of determination; AUPRC: area under the precision–recall curve; MCC: Matthews correlation coefficient; FID: frchet inception distance; ROC: receiver operating characteristic; AUC: area under the curve; FLOPs: floating-point operations.

Results from general bearing benchmarks mainly demonstrate algorithmic feasibility under controlled conditions and cannot fully represent performance in operational wind farms. Wind-turbine-specific validation should additionally consider variable loads, environmental disturbances, sensor heterogeneity, temporal drift, and the scarcity of verified fault labels. Therefore, the datasets and metrics listed in [Table 1](#) should be selected according to the intended diagnostic task and deployment environment rather than applied as a uniform evaluation standard. Future studies should therefore establish standardized data partitions and cross-condition validation protocols to improve the comparability of reported results. In addition, uncertainty intervals and repeated-run statistics should be reported whenever possible to distinguish genuine model improvements from random variation.

3. DATA AUGMENTATION AND MULTI-SOURCE FUSION

This section examines three data-level strategies for wind turbine drivetrain diagnosis: generative augmentation for scarce fault samples, feature-space governance under noisy and non-stationary conditions, and multimodal fusion of heterogeneous monitoring data.

3.1. Generative sample augmentation

In the face of the data dilemma characterized by fault sample scarcity and extreme class imbalance in wind turbine drivetrain components, GANs, variational autoencoders (VAEs), and their derivative variants combined with diffusion models have emerged as crucial technical pathways for synthesizing high-fidelity fault samples and bridging data gaps. The core logic of these methods is to learn the majority-class data distribution and then generate minority-class anomaly samples through perturbation or mapping, thereby reconstructing a balanced class ratio within the training set. Synthetic-data quality and diagnostic performance should be evaluated separately. Accuracy, precision, recall, and F1-score evaluate the task performance of the diagnostic model after augmentation. For synthetic images, frchet inception distance (FID) and related image distribution metrics can assess discrepancies in image feature distributions. Such

Table 2. Quantitative comparison of generative augmentation methods

Method	Dataset	Reported performance
AFDVGAN	Electric-locomotive and aerospace-bearing data	99.81% and 99.16% diagnostic accuracy ^[35]
AVAEGAN-ADS	HENU; CWRU; NASA bearing datasets	97.14%, 97.52%, and 98.17% diagnostic accuracy ^[36]
PIFE-DM	Simulated; CWRU; bearing test-rig data	97.00% and 96.33% accuracy on two real datasets ^[37]

image-based metrics cannot directly establish the physical plausibility or fault-feature fidelity of generated diagnostic signals; signal-domain validity therefore requires separate verification.

Conditional Wasserstein GANs incorporate class labels and Wasserstein distance to stabilize adversarial training, suppress mode collapse, and improve recognition of scarce bearing-fault samples under imbalanced conditions^[28]. Under imbalanced few-shot scenarios, gradient-penalized Wasserstein GANs generate high-quality fault signals while improving convergence stability and classification performance^[29].

Hierarchical VAE–diffusion models learn latent mappings from healthy to damaged states to augment minority-class samples. With only 1%–2% real fault samples, the hybrid framework achieves a reconstruction mean squared error (MSE) below 0.005 and a small FID deviation, thereby reducing classifier overfitting and improving damage-recognition stability under unseen sea states^[30]. Comparatively, GANs, VAEs, and diffusion models exhibit different performance characteristics because of their distinct distribution-learning mechanisms. GAN-based methods directly approximate real fault distributions through adversarial training and therefore tend to generate sharper and more fault-sensitive samples, but their minimax optimization is susceptible to unstable convergence and mode collapse^[31]. VAE-based methods provide more stable training and continuous latent representations through probabilistic encoding, although variational regularization may over-smooth transient impacts and weak fault signatures^[32]. Diffusion models generally achieve a better balance between sample fidelity and diversity through iterative denoising, but their multi-step generation process introduces substantially higher computational costs^[33]. Consequently, GANs are more suitable for rapid sample augmentation, VAEs are preferable for stable latent-space modeling, and diffusion models are advantageous when generation quality is prioritized and sufficient computing resources are available.

However, their effectiveness depends strongly on the completeness of healthy-data distributions, while extremely limited minority samples may still induce overfitting. In addition, iterative diffusion training and sampling can substantially increase computational cost^[34]. To provide additional quantitative evidence beyond the studies discussed above, Table 2 compares three newly included generative augmentation methods by benchmark dataset, diagnostic performance, and computational characteristics.

The studies further demonstrate that adversarial models can provide highly accurate augmentation under severe class imbalance, whereas VAE–GAN hybrids offer more stable latent representation learning. Diffusion-based augmentation achieves competitive diagnostic performance but generally incurs greater sampling costs because of its iterative generation process.

3.2. Structural governance of feature spaces

Wind turbine monitoring data are frequently affected by noise, missing values, and operating-condition fluctuations, which destabilize feature distributions and reduce diagnostic reliability. The core objective of feature space structural governance is to rectify feature distribution shifts through targeted interventions and reconstruct the orderliness of the feature space, thereby establishing a low-entropy, highly separable feature foundation for subsequent multimodal information fusion. This framework encompasses a comprehensive

technical pipeline spanning from low-level signal enhancement and spatial geometric structure optimization to data reconstruction under complex operating conditions.

Situated at the lowest level of the data processing pipeline, feature engineering and signal enhancement technologies serve as the cornerstone for constructing a robust feature space. For non-stationary machinery signals, time-frequency analysis can reveal evolving fault-related components and provide more interpretable inputs for subsequent models^[38].

Manifold learning maps high-dimensional disordered features into structured low-dimensional representations. Spectral dictionary and feature-cloud methods use signal decomposition and reconstruction to distinguish different health states in a latent manifold^[39]. Recent supervised manifold-learning methods further combine adaptive neighborhood construction with discriminative multi-feature fusion to preserve local geometric relationships while explicitly incorporating fault-class information, thereby improving the separability of health-state representations under complex operating conditions^[40]. This structured spatial mapping significantly enhances the stability and generalization capability of diagnostic models under few-shot conditions, effectively counteracting the distribution shifts induced by operational fluctuations. Nevertheless, manifold-based feature governance assumes that samples with similar degradation states preserve local geometric consistency in the latent space. Therefore, it is more effective for gradual degradation processes and condition monitoring tasks. For abrupt failures, severe sensor contamination, or strong cross-domain shifts, the learned manifold structure may become distorted, requiring complementary strategies such as robust representation learning or domain adaptation.

In extreme scenarios lacking fault data, feature-space governance becomes the precise delimitation of the normal domain boundary. For zero-fault-sample scenarios, generating out-of-distribution samples from normal monitoring data and constructing highly discriminative feature pairs within the feature space compels the model to learn the decision boundary between normal and anomalous classes. Essentially, this approach reinforces the feature-space topology by artificially introducing contrastive constraints, thereby enabling effective anomaly detection. Beyond synthetic boundary construction, distance-aware out-of-distribution detection can combine spectral normalization and Gaussian-process uncertainty estimation to distinguish known gearbox health states from previously unseen faults, thereby reducing the risk of overconfident closed-set decisions^[41].

In summary, physics-prior guidance, manifold mapping, and out-of-distribution boundary construction jointly improve feature separability and robustness, providing reliable representations for subsequent multimodal diagnosis.

3.3. Cross-modal physical fusion strategies

Multimodal fusion requires spatio-temporal integrity and consistency across modalities. Addressing data loss caused by sensor failures, conventional interpolation methods fail to restore the inherent underlying dynamics of the signals. To address these gaps, physical-constraint-based generative imputation models can reconstruct missing wind-turbine SCADA values while enforcing consistency with turbine operating characteristics^[42].

For vibration-based diagnosis, the core challenge is to effectively integrate one-dimensional signal features and two-dimensional time-frequency representations within the feature space. Advanced integration schemes, such as the two-stream feature fusion residual network (TSFFResNet), align one-dimensional vibration features with two-dimensional time-frequency representations across parallel streams. This enables bearing fault identification to maintain accuracy above 99% even under dynamic loading conditions^[43]. However, vibration signals are typically sampled at the kilohertz level, whereas SCADA variables are

Table 3. Summary of data processing and multimodal fusion technologies for wind turbine intelligent O&M

Technical paradigm	Representative methods	Core engineering value	Potential deployment	System integration challenges
Multi-sensor fusion	AcvGraph	Fuses acoustic and vibration data through graph-based correlation learning ^[48]	Edge/station	Graph construction and sensor alignment
Non-stationary analysis	Spectral correlation, Wigner-Ville, TAR	Models cyclostationary and non-stationary turbine vibration ^[49]	Station/cloud	Model order and operating-state variability
SCADA-vibration fusion	Contrastive learning, LPC features	Combines low-frequency SCADA with high-frequency vibration for fault detection ^[50]	Edge/station	Sampling-rate alignment and synchronization

O&M: Operation and maintenance; TAR: true acceptance rate; SCADA: supervisory control and data acquisition; LPC: linear predictive coding.

recorded much less frequently, for example at 1 Hz or at 10-min intervals, making direct point-to-point fusion prone to temporal mismatch and information distortion^[44]. A complementary main-drive-chain diagnosis study jointly extracted fault features from SCADA and vibration-monitoring data and evaluated the data-fusion model using actual wind-turbine cases^[45]. Existing deep fusion frameworks mainly address this problem through window-level aggregation and resampling, multi-rate dual-stream encoders combined with cross-attention or hierarchical temporal pooling, and event- or operating-condition-aware alignment based on rotational-speed intervals, alarm events, or state transitions. SCADA data provide long-term operational information, including power, temperature, and control variables, making them suitable for system-level health assessment. However, their relatively low sampling frequency limits sensitivity to early mechanical degradation. In contrast, vibration signals contain richer fault-related high-frequency information but are more vulnerable to sensor installation conditions and environmental noise^[46]. Therefore, SCADA-dominated fusion is preferable for fleet-level anomaly monitoring, whereas vibration-centered fusion is more appropriate for early-stage drivetrain fault diagnosis.

The ultimate form of multimodal fusion involves mapping the data space back to the physical entity space to achieve interpretable proactive warnings. By combining operational data with a digital representation of the turbine, the system can support real-time condition monitoring, transferable diagnosis, and earlier fault detection^[47].

In summary, multimodal diagnosis requires temporal reconstruction, multi-rate alignment, and physics-informed integration to combine transient vibration signatures with long-term SCADA operating information. Table 3 compares the corresponding methods, deployment locations, and implementation challenges. Such complementary fusion of transient and long-term information not only mitigates the information loss caused by sampling-rate mismatch but also enhances early-fault sensitivity and diagnostic robustness under varying operating conditions. Nevertheless, the reliability and generalization of these multimodal schemes still depend heavily on data synchronization, labeling effort, and the interpretability of the fused decision process.

4. ADVANCED LEARNING PARADIGMS FOR COMPLEX WIND POWER DATA

Building upon the data challenges discussed in Section 3, this section shifts focus toward constructing efficient learning paradigms at the algorithmic level. To address the dual constraints of sample scarcity and operational fluctuations, we delineate a technical roadmap integrating semi-supervised and contrastive learning, cross-domain transfer learning, and adaptive continual learning, following the emerging paradigm of lifelong adaptive learning under non-stationary environments^[51]. This integrated approach provides a robust pathway for high-precision, life-cycle diagnosis of wind turbine drivetrain components under complex operating conditions.

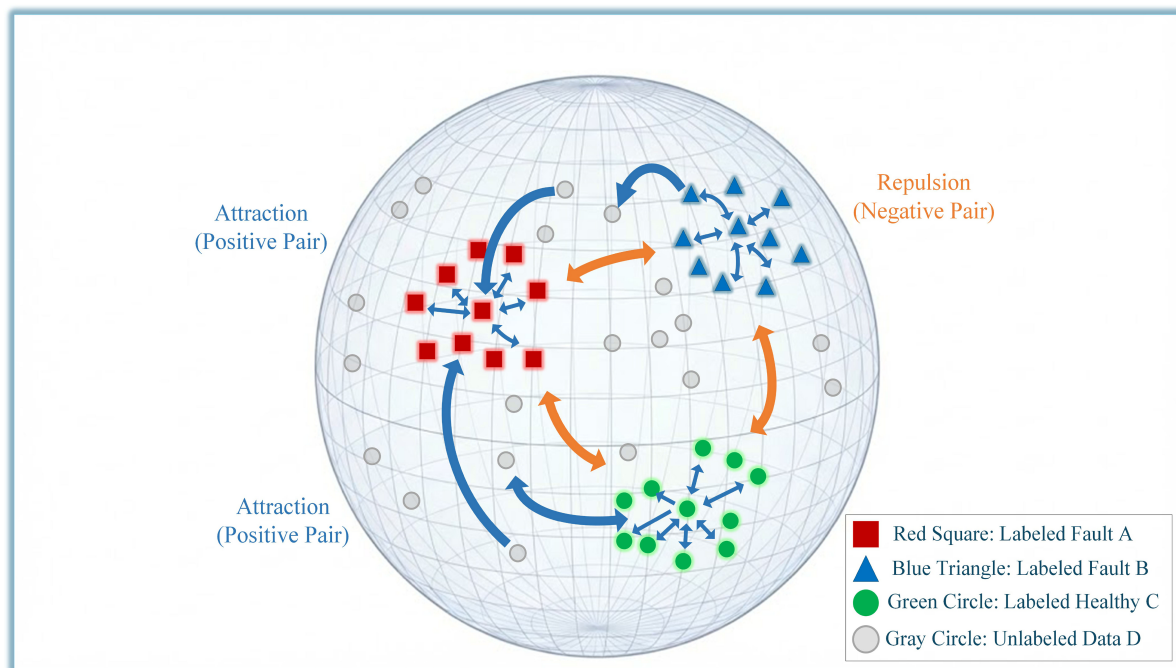


Figure 4. Feature space optimization in semi-supervised contrastive learning.

4.1. Semi-supervised and contrastive learning under label scarcity

Leveraging massive unlabeled data to optimize the feature-space distribution is pivotal for improving data quality. Distinct from the generative models in Section 3, semi-supervised and contrastive learning focus on mining the intrinsic topological structures of the data. For instance, the unified imbalanced semi-supervised contrastive learning framework integrates supervised and unsupervised contrastive losses to effectively rectify decision boundaries. As illustrated in Figure 4, supervised contrastive learning pulls same-class representations together while separating different classes in the embedding space^[52]. By introducing consistency regularization and pseudo-labeling strategies, semi-supervised learning methods such as FixMatch can exploit unlabeled data distributions and improve decision-boundary robustness under limited labeled conditions^[53]. Wind-turbine-specific studies have further extended this paradigm through matching contrastive learning for fault diagnosis with imbalanced SCADA data^[54]. These results indicate that diagnostic performance depends not only on the quantity of unlabeled data but also on class-aware pair construction, identification of easily confused samples, and preservation of inter-class margins. This mechanism effectively mitigates the risk of overfitting caused by label scarcity, enabling the model to construct robust discriminative boundaries even with minimal labeled samples. Despite these advantages, semi-supervised and contrastive learning methods are highly dependent on the assumption that unlabeled samples share similar distributions with labeled data. Under severe operating condition changes, inaccurate pseudo-label propagation may amplify classification errors. Therefore, these methods are particularly suitable for wind turbines with abundant unlabeled operational data and limited labeled fault samples, whereas transfer learning or domain adaptation should be preferred when large distribution discrepancies exist.

4.2. Transfer learning for adaptive working conditions and diverse scenarios

When a model is required to generalize from one wind farm to another with a vastly different geographical environment, or from steady operating conditions to extreme gust conditions, the significant shifts in data distribution often lead to model failure. Transfer learning provides a systematic solution to these cross-domain diagnostic challenges by mining the invariant features shared between the source and target domains^[55].

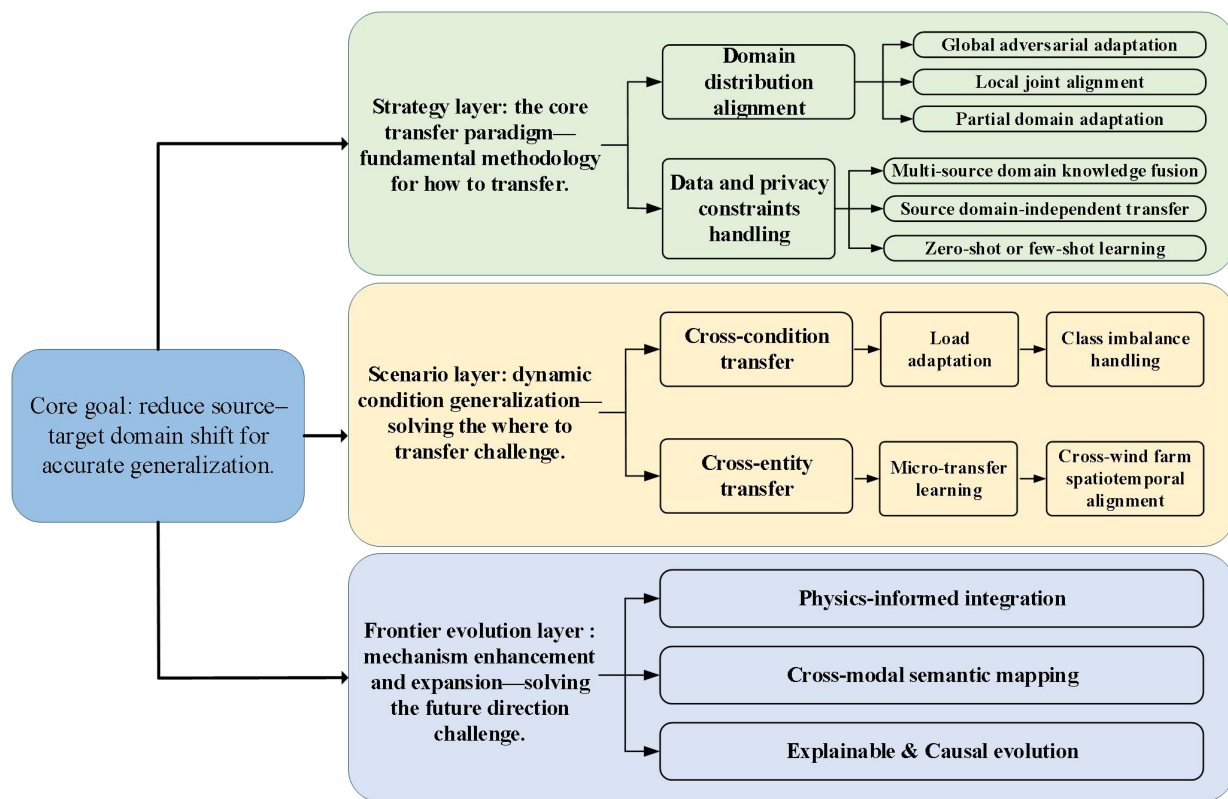


Figure 5. Taxonomy of transfer learning for cross-domain fault diagnosis in wind turbines.

To systematically categorize existing cross-domain diagnostic methodologies, this paper adopts a hierarchical transfer learning framework, as illustrated in Figure 5. From the three perspectives of methodology, application scenarios, and frontier evolution, current research is organized into three layers: the strategy layer, the scenario layer, and the frontier evolution layer. The strategy layer focuses on the core algorithmic question of “how to transfer”, encompassing distribution adaptation techniques ranging from global adversarial learning to fine-grained conditional alignment. The scenario layer addresses the practical engineering question of “where to transfer”, with particular emphasis on cross-condition load variations and spatio-temporal discrepancies among different wind turbines and wind farms. The frontier evolution layer focuses on physics-informed integration, cross-modal semantic mapping, and explainable and causal transfer learning, thereby outlining the future development trajectory of cross-domain fault diagnosis.

Following this three-layer framework, this section provides an in-depth analysis of key methodologies, application scenarios, and frontier developments in cross-domain wind turbine fault diagnosis.

4.2.1. Core transfer paradigms

To directly address the distribution discrepancies between source-domain data and target-domain operating conditions, adversarial domain adaptation has emerged as a fundamental technical route for cross-turbine diagnosis. Represented by the domain-adversarial neural network, the core philosophy of global adversarial adaptation stems from the zero-sum game theory of GANs.

Global adversarial adaptation uses a domain discriminator and gradient reversal layer to reduce marginal-distribution shifts between source and target domains. However, because it does not explicitly preserve

class-conditional structures, it may blur fault boundaries and induce negative transfer when fault proportions or label spaces differ across wind farms.

Conditional subdomain adaptation instead aligns corresponding fault-related subdomains using class-aware discriminators, fault prototypes, joint feature-label distributions, or confidence-weighted pseudo-labels. By reducing intra-class variation while preserving inter-class margins, it is better suited to multi-fault, variable-speed, and noisy conditions.

Nevertheless, fine-grained alignment is not universally superior. Its performance depends heavily on the reliability of target-domain pseudo-labels and class-prototype estimation; incorrect pseudo-labels, severe class imbalance, or absent target-domain fault classes may cause erroneous local matching and amplify negative transfer. It also introduces additional computational and hyperparameter-tuning costs compared with a single global discriminator. Therefore, global alignment remains suitable for scenarios dominated by coarse operating-condition shifts with relatively consistent class structures, whereas conditional subdomain alignment is more appropriate for multi-fault and strongly nonstationary wind power scenarios, provided that confidence estimation, robust sample selection, and class-imbalance-aware mechanisms are incorporated. Global-local adaptation networks combine both mechanisms and have demonstrated effectiveness in pitch-bearing diagnosis^[56].

When source and target domains have inconsistent label spaces, partial domain adaptation uses attention-based weighting to suppress source-only classes and reduce negative transfer in gearbox diagnosis^[57]. A graph causal framework with adaptive expert ensembles provides a distinct approach to addressing distributional shifts in fault diagnosis^[58]. Adversarial adaptation has further been extended from vibration analysis to unmanned aerial vehicle (UAV)-based visual inspection through soft-mask-guided Faster R-CNN models^[59].

In real-world wind power scenarios, diagnostic systems often encounter data from multiple turbines or heterogeneous operating conditions. Multi-source domain transfer and meta-learning frameworks have significantly expanded the generalization boundaries by systematically integrating knowledge from diverse domains. For instance, a multi-domain multi-task learning framework leverages data from various experimental test rigs to construct models for deployment on real-world turbines^[60]. This approach, combined with models such as CNN-LSTM and improved GANs, enables comprehensive performance evaluation within a unified framework^[61]. Simultaneously, source-free unsupervised domain adaptation methods have demonstrated exceptional adaptability in large-scale experiments involving multiple wind farms and diverse turbine units^[62,63].

Standard transfer learning typically assumes that data from both the source and target domains are simultaneously accessible. However, in real-world wind power scenarios, commercial privacy protection and limited bandwidth for massive data transmission often constrain access to raw fault samples. To address these data silos and privacy challenges, source-free transfer strategies have emerged. For instance, source-free domain adaptation under partial information has been investigated for rotating-machinery fault diagnosis without access to the original source-domain data during adaptation^[64]. This offers a relevant methodological basis for settings in which sharing raw source data is restricted, but does not by itself establish cross-wind-farm performance.

Furthermore, considering the time-varying characteristics throughout the full life cycle of a device, a micro-transfer learning mechanism achieves more streamlined, unit-level knowledge transfer by modeling multiple differentiated distributions^[65]. Regarding multi-source heterogeneous operating conditions, domain adaptation region mechanisms facilitate matching of both global and category-wise distributions by

constructing adaptive intermediate distributions^[66].

Beyond data access constraints, another extreme challenge arises when the target domain encounters previously unseen fault types, as encountered in zero-shot or few-shot scenarios. In response to such data deserts, single-domain generalization methods offer a viable solution for online diagnosis without target data by learning features that remain invariant to future domain shifts^[67]. Under few-shot conditions, strategies that combine the transformation of 1D vibration signals into 2D images with deep residual networks have proven effective. For instance, reconstructing feature representations using Gramian angular difference fields has shown high reliability for imbalanced and heterogeneous data^[68,69]. To mitigate classification bias from incomplete labeling, entropy-weighted manifold alignment dynamically adjusts sample weights^[70].

4.2.2. Generalization across dynamic operating conditions

Cross-condition diagnosis requires models to learn operating-condition-invariant features. Contrastive ensemble transfer learning uses stochastic contrastive regularization^[71]. Concurrently, the cross-machine deep subdomain adaptation network demonstrates superior performance under variable-speed conditions based on the measurement of subdomain distribution shifts^[72].

To address sample scarcity, the conditional VAE-GAN synthesizes fault data for missing operating conditions^[73]. Furthermore, methods such as the class-imbalance-aware deep adversarial adaptation network and semi-supervised weighted pseudo-labeling effectively resolve the concurrent engineering challenges of condition shifts and class imbalance. They achieve this by optimizing the imbalanced feature space and introducing pseudo-labeling mechanisms, respectively^[74,75].

At the level of large-scale deployment, cross-wind-farm transfer must bridge the generalization gap induced by geographic and microclimate disparities. Research has consequently shifted from single-turbine diagnosis to wind-farm-level modeling. For instance, through structural optimization, models such as the multi-layer bidirectional LSTM network achieve unified diagnosis for multi-class blade faults across an entire wind farm^[76]. Furthermore, integrating the maximum mean discrepancy algorithm into the alignment of SCADA spatio-temporal features has significantly enhanced cross-turbine generalization capabilities^[77].

Simultaneously, to address sample scarcity and limited computational resources for individual turbines, researchers have adopted a micro-transfer learning paradigm. This approach utilizes pre-trained CNNs to extract universal feature weights, followed by parameter fine-tuning exclusively on specific layers. By retaining the generalized knowledge from the source domain while enabling rapid adaptation to specific entities, this strategy provides a highly flexible pathway for addressing the complex and highly variable main shaft bearing faults across different wind farms^[78].

4.2.3. Physics-informed enhancement and frontier technology integration

Traditional transfer learning can overlook physical consistency across wind farms. Physics-informed machine learning offers a general route for combining data with governing knowledge^[79], while non-stationary wind-turbine diagnosis can also exploit physically meaningful, speed-varying fault-frequency trajectories through generalized demodulation^[80]. These studies demonstrate that physics-informed enhancement can reduce reliance on purely statistical correlations, although its effectiveness remains dependent on the accuracy and cross-turbine validity of the embedded physical priors.

With the widespread adoption of UAV-based visual inspection, adapting diagnostic models to rare or previously unseen blade defects has become increasingly important. Knowledge-augmented vision-language models combine drone-captured blade images with retrieved defect descriptions and expert knowledge,

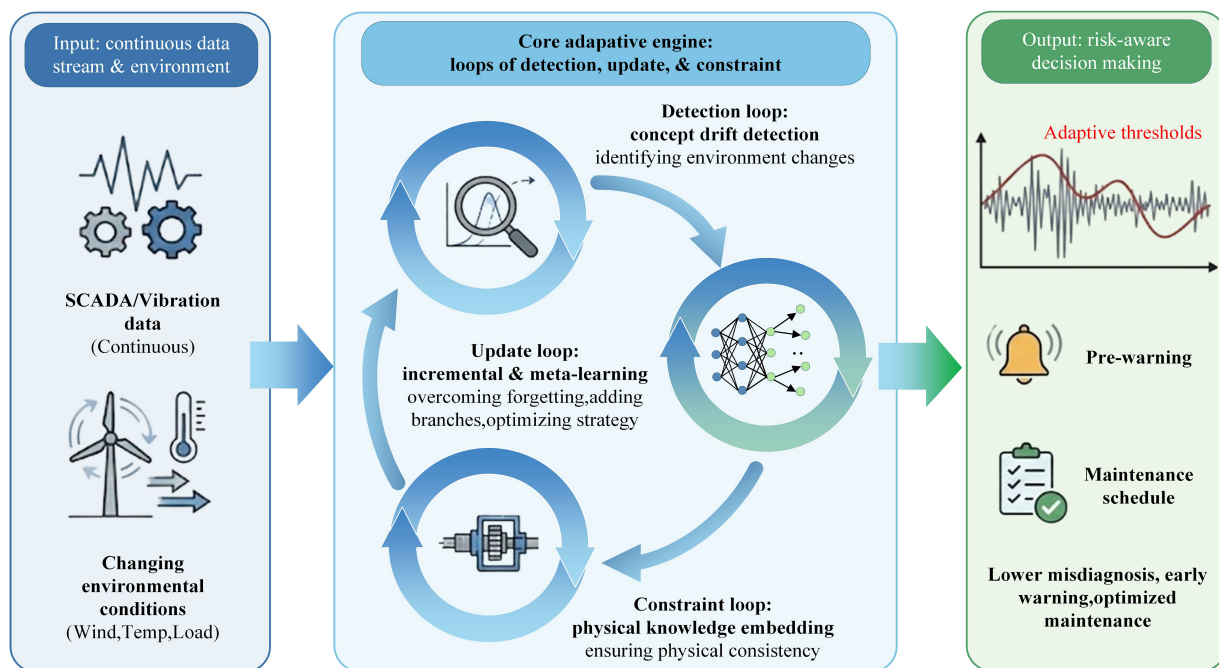


Figure 6. Physics-constrained incremental learning architecture for wind turbine O&M. O&M: Operation and maintenance; SCADA: supervisory control and data acquisition.

enabling data-efficient and interpretable classification, localization, and severity assessment for wind-turbine blade inspection^[81].

Furthermore, conventional transfer models rely heavily on feature correlations and often lack logical explanations for fault root causes. Transfer strategies incorporating causal inference aim to disentangle spurious correlations induced by environmental fluctuations from intrinsic causalities driven by component failures. Moving forward, interpretable transfer learning will not merely alert O&M personnel to a fault, but will also transparently explain why knowledge learned under one operating condition remains applicable under an unseen condition^[82]. This paradigm shift will fundamentally enhance the trustworthiness of automated decision-making within expert systems.

4.3. Adaptive and intelligent decision architectures

As wind turbines progress towards large-scale and deep-offshore deployments, the stochasticity of operating environments and complex working conditions impose increasingly stringent demands on the real-time responsiveness, robustness, and physical consistency of diagnostic models. To enable maintenance systems to evolve continuously in open environments, the research community is increasingly adopting continual-learning strategies that update models while limiting catastrophic forgetting^[83].

Figure 6 illustrates a physics-constrained incremental learning architecture tailored for wind turbine O&M. This framework constructs a closed-loop system comprising a detection loop, an update loop, and a constraint loop, ultimately yielding risk-aware dynamic decision-making outputs. This section systematically discusses cutting-edge advancements in this framework across five dimensions: learning mechanisms, online evolution, policy optimization, underlying architecture, and decision assurance.

4.3.1. *Continual learning paradigms to mitigate catastrophic forgetting*

A fundamental challenge in deploying data-driven models for wind turbine health management lies in continuously adapting to novel fault patterns without erasing previously acquired knowledge—a phenomenon widely recognized as catastrophic forgetting^[84]. To address this challenge, the research community has proposed various incremental and lifelong learning frameworks, aiming to strike an optimal balance between stability and plasticity through diverse architectural designs and algorithmic mechanisms.

Among these, a representative paradigm leverages pseudo-labeling mechanisms to filter high-confidence samples from newly emerging unlabeled data streams. This approach augments the training set, thereby achieving adaptive updates with minimal manual annotation costs. For instance, uncertainty-aware pseudo-label selection can incorporate confident unlabeled bearing samples into semi-supervised transfer diagnosis under data imbalance^[85]. While this strategy is attractive in annotation-scarce scenarios, its efficacy depends on probability calibration, semi-supervised learning assumptions, and distinguishing aleatoric from epistemic uncertainty^[86-88].

Another category of approaches adheres to the philosophy of dynamic architectural expansion. By incrementally appending new branches to the network to accommodate novel fault classes, these methods explicitly preserve previously acquired knowledge at the structural level. Notable examples include the partial matching segmented CNN incremental learning framework, which continually extracts new fault features from data streams while maintaining a diagnostic accuracy exceeding 98% under varying operating conditions^[89]. Similarly, inverted transformer-based lifelong learning can add and prune branches for emerging fault classes, while regularization-based continual learning provides a complementary mechanism for retaining earlier knowledge^[90,91].

Although this branch-expansion strategy is intuitive for retaining historical knowledge and achieves exceptional accuracy in diagnosing unknown faults^[92], it comes with an inherent trade-off: the model architecture inevitably bloats as the number of tasks increases. Consequently, this raises significant concerns regarding computational and storage efficiency during long-term deployment, particularly on resource-constrained edge devices.

To mitigate these inefficiencies, exemplar-free replay approaches eliminate the need to store historical data samples. The adaptive prototype correction and separation network exemplifies this trajectory. It utilizes optimal transport theory to calibrate historical class prototypes within a newly updated feature space. By integrating contrastive learning to enhance feature separability, this method achieves exceptional diagnostic accuracies of 99.01% and 97.36% on benchmark datasets, all without retaining any legacy samples^[93]. Although this strategy alleviates data privacy concerns and storage burdens, it shifts the engineering challenge to maintaining the long-term robustness and drift resistance of these prototype representations.

Furthermore, a subset of studies integrates transfer learning principles into incremental frameworks, leveraging historical experience to accelerate the learning process. For instance, the trend-constrained incremental transfer prognosis approach uses trend constraints to guide prognostic updates under distribution drift, enhancing both the stability and efficiency of incremental predictions^[94].

Overall, spanning from pseudo-label augmentation and architectural expansion to prototype correction and transfer acceleration, these frameworks collectively constitute a multi-dimensional toolbox for overcoming catastrophic forgetting. However, practical constraints often limit their efficacy, indicating the absence of a universally optimal solution. Consequently, the selected strategy must be precisely tailored to the unique demands of the deployment scenario.

4.3.2. *Online learning and dynamic model updating*

As data-driven models transition into online deployment, they must handle continuous data streams, distribution drift, and time-varying noise. Practical feasibility therefore depends on concept-drift adaptation and real-time noise-robust updating^[95,96]. Under such scenarios, the theoretical capabilities of incremental learning must be operationalized into executable dynamic updating mechanisms, thereby sustaining the timeliness and precision of the prognostic outcomes.

To reduce the computational burden of large-scale industrial data streams, a key strategy is intelligent sampling, which extracts a representative subset from event logs by quantifying the significance of data trajectories. For instance, when a sampling-based next-event prediction method is applied to wind turbine maintenance logs in conjunction with an LSTM network, utilizing a 30% sampling ratio yields a 3.631-fold improvement in prediction efficiency and a concurrent 6.896% increase in accuracy^[97]. However, sampling strategies inherently carry the risk of information loss or the introduction of statistical bias. Consequently, researchers must carefully balance gains in processing speed with the preservation of information fidelity.

Simultaneously, edge-based incremental learning can support online evolution by updating models directly on inference devices. Because wind-turbine data are affected by sensor noise and uncertainty, online updating should also incorporate noise-adaptation capabilities.

Ultimately, the engineering realization of lifelong online capabilities hinges on the synergy of three pivotal components: efficient data stream management, edge-based incremental update protocols, and embedded online filtering with noise adaptation mechanisms. Collectively, these elements forge the technological bedrock required to support autonomous, real-time adaptation.

4.3.3. *Meta-learning and RL-driven adaptive strategy optimization*

In contrast to passive, reactive updating in response to continuous data streams, the research frontier increasingly leverages meta-learning and reinforcement learning (RL) to shift adaptation from passive updating to proactive optimization. This advancement gives models the intrinsic ability to autonomously optimize their learning strategies and decision-making processes, enabling them to navigate dynamic environments with unprecedented intelligence. Ultimately, this paradigm shift aims to alleviate the heavy reliance on manual task design and meticulous policy tuning typically required in complex, multi-objective scenarios.

The profound value of meta-learning lies in its capacity to automate the learning process itself. For instance, by automatically generating auxiliary task labels via a meta-auxiliary generation network, supplementary degradation information can be provided for RUL estimation, thereby augmenting the primary task performance^[98]. Concurrently, to address the prevalent issue of inaccurate labels in real-world data, certain meta-learning approaches dynamically aggregate online soft labels based on cross-task feature similarities. This mechanism guides model training and substantially elevates diagnostic accuracy within few-shot and highly noisy environments^[99].

While such methodologies significantly curtail manual engineering efforts and bolster robustness, their efficacy remains highly contingent upon the quality of the initial meta-knowledge. Furthermore, continuously updating the meta-learner introduces additional computational overhead.

In contrast to meta-learning's emphasis on learning to learn^[100], RL provides a direct pathway for optimizing high-level O&M policies. RL is capable of achieving multi-objective optimization through dynamic trade-offs

- for example, simultaneously managing the risk of blade fatigue failure while minimizing both maintenance costs and carbon emissions^[101].

Furthermore, when integrated with formal modeling tools such as Petri nets, RL can navigate complex state-action spaces to learn optimal condition-based maintenance policies. This synergy can yield exceptional outcomes, such as attaining a 99.4% system availability alongside minimized operational expenditures^[102]. However, effective RL application often relies on high-fidelity environmental simulations and complex reward-function engineering. Moreover, its inherently low sample efficiency can pose a significant barrier to deployment within safety-critical real-time systems.

In essence, the integration of meta-learning and RL signifies a profound leap in wind turbine diagnostics: transitioning from purely data-driven paradigms to cognitive-driven intelligence. By endowing models with the learning-to-learn capability, meta-learning facilitates rapid adaptation; conversely, RL optimizes O&M policies through the simulation of sequential decision-making processes. This evolution empowers the system to transition from passive data reception to proactive environmental adaptation, ultimately striking an optimal trade-off between diagnostic reliability and operational expenditures under highly uncertain operating conditions.

4.3.4. Adaptive model architectures and feature learning

Beyond external learning strategies, the resilience and adaptability of a model within highly dynamic and complex environments depend fundamentally on the intrinsic capacity of its underlying architecture to encode and process spatio-temporal dynamics and physical constraints^[103-105]. This perspective underscores the need to move beyond merely analyzing how a model learns to investigate the endogenous robustness of its architectural design. Ultimately, such intrinsic structural integrity endows the model with an inherent cross-domain generalization capability when confronting complex and highly variable operating conditions.

A primary architectural innovation involves explicitly modeling the relational structures within multi-sensor systems. For instance, the recurrent graph convolutional network with uncertainty estimation (RGCNU) learns spatial and temporal dependencies in condition-monitoring data. Its uncertainty model assumes Gaussian-distributed RUL predictions and estimates a predictive mean and variance using a likelihood-based loss, rather than quantile regression^[106]. However, the efficacy of graph-based models is heavily reliant on the accuracy of the constructed graph topology, a task that proves exceedingly difficult in highly non-stationary scenarios.

Complementary to purely data-driven relational modeling is anchoring the representation space by integrating domain-specific physical knowledge. Physics-informed architectures, for example, can combine physics-based fatigue and grease-degradation kernels with data-driven neural networks, improving physical consistency and uncertainty-aware long-term degradation estimation for wind turbine main bearings^[107]. More broadly, physics-informed machine learning combines observational data with scientific knowledge to improve consistency and data efficiency, but its reliability still depends on the validity of the embedded assumptions^[108]. Nevertheless, the effectiveness of such methodologies is inherently bounded by the accuracy and completeness of the integrated physical models, which may ultimately struggle to encompass the full spectrum of potential failure modes.

To achieve broad robustness without relying on specific physical laws, fusion-based ensemble learning frameworks provide an alternative architectural pathway. By constructing hybrid models that encompass diverse algorithmic components coupled with feature optimization, this class of methods demonstrates exceptional generalization capabilities across varying operating conditions. They can achieve high R^2 values

while significantly reducing prediction errors^[109]. The associated cost, however, is a surge in model complexity. Furthermore, when compared to concise single models, these ensemble approaches often confront substantial challenges regarding interpretability and deployment maintenance.

Consequently, spanning from graph relational modeling and physics-constraint fusion to robust ensemble architectures, these advancements collectively constitute the adaptive infrastructure at the structural level of the model. Ultimately, this infrastructure endows the system with the indispensable capabilities for perception, representation, and reasoning within the dynamic, complex environments of real-world wind turbines^[110].

4.3.5. Adaptive thresholds and reliable decision-making under concept drift

The ultimate operationalization of a model's adaptive capabilities manifests in its decision boundaries and alarm mechanisms: adaptive thresholds and concept drift detection constitute the ultimate safeguard for ensuring decision reliability. The core paradigm shift lies in transitioning from static, predefined thresholds to dynamically adjusted mechanisms that evolve concurrently with model confidence and data distributions, thereby effectively mitigating both false alarms and missed detections.

Probabilistic forecasting frameworks provide critical structural support for this transition. Through rigorous uncertainty quantification, decision-making can be translated into risk-informed scheduling based on confidence intervals. For instance, an RUL prediction model for gearbox pump failures, underpinned by a Bayesian neural network, reported that the true failure time fell within a range of approximately ± 5.3 h of the predicted value with a 97.5% confidence level. Statistically, this furnishes a robust foundation for maintenance scheduling^[111]. Crucially, these probabilistic outputs can be directly used to formulate and continuously adjust dynamic alarm thresholds.

Furthermore, specific research endeavors optimize the decision-making process for hard-to-classify samples: misclassifications are most prone to occur when a sample's probabilities of belonging to the positive and negative classes are nearly equal. To address this vulnerability, purpose-built SVM ensemble models are deployed to identify and improve the classification accuracy of these ambiguous instances, thereby substantially bolstering overall decision reliability^[112].

To sustain long-term reliability amidst slowly drifting operational baselines, a possible design is to combine normal behavior models with residual-based alarm thresholds. Threshold selection and recalibration should be evaluated against false-alarm rates and fault-detection delays under the intended operating conditions; no single residual quantile should be treated as universally appropriate.

However, unconstrained threshold adaptation may absorb slowly evolving fault signatures into the updated normal baseline, thereby delaying alarms for genuine component degradation. To distinguish benign environmental or operational drift from irreversible health deterioration, current methods increasingly employ condition-conditioned normal behavior models and dual-timescale updating mechanisms: fast adaptation compensates for variations in wind speed, load, ambient temperature, and seasonal conditions, whereas a slowly updated or fixed reference baseline preserves long-term degradation information. Threshold recalibration can also be gated by residual persistence, monotonic degradation indicators, change-point detection, and physics-informed health constraints, and should be suspended when deviations exhibit sustained, component-specific, or irreversible patterns. In safety-critical applications, adaptive thresholds should therefore be combined with conservative alarm boundaries and independent degradation detectors to prevent baseline updating from masking incipient faults.

Table 4. Learning strategies for wind turbine predictive maintenance

Technical paradigm	Representative technologies	Core advantages	Key engineering metrics	Applicable scenarios
Incremental learning	Incremental-learning fault diagnosis taxonomy	Mitigates catastrophic forgetting across evolving fault classes	Comprehensive review of dynamic-system ILFD methods ^[113]	Continuous integration of novel fault classes
Online learning	Lightweight 1D-CNN-Transformer	Real-time, noise-robust edge diagnosis	1.98 ms single-sample latency on Raspberry Pi 4B ^[114]	Edge computing and real-time condition monitoring
Meta, RL	CNTE-MAML	Few-shot adaptation on simulated and field wind-turbine data	Outperforms seven comparison methods under limited samples ^[115]	Few-shot, high-noise environments, and complex policy optimization
Adaptive architecture	FFT + GraphSAGE + GATConv	Models inter-sensor correlations with automated optimization	99.73% precision on NREL GRC data ^[116]	Cross-condition transfer and non-stationary operations
Adaptive decision-making	Continual test-time domain adaptation	Adapts to covariate and label shifts in data streams	3.78% average diagnostic-accuracy improvement ^[117]	Continual adaptation under evolving data streams

ILFD: Incremental learning-based fault diagnosis; 1D-CNN: one-dimensional convolutional neural network; RL: reinforcement learning; FFT: fast Fourier transform.

These adaptive detection mechanisms underscore the industry’s shift toward more resilient and autonomous diagnostic systems. However, the diversity of operational environments and data characteristics necessitates a broader evaluation of available methodologies beyond individual frameworks. From a practical standpoint, the evaluation of such adaptive detection schemes should therefore go beyond isolated accuracy metrics and incorporate robustness across diverse wind regimes, alarm latency, false-alarm rates, and the ability to preserve detection sensitivity under gradual degradation.

Table 4 provides a comprehensive summary of advanced learning strategies for wind turbine predictive maintenance. This encompasses representative methodologies across various technological paradigms, their core advantages, and their practical engineering efficacy. The table not only encapsulates the applicable scenarios and performance metrics of these diverse methods under real-world operating conditions but also serves as a critical reference guide for selecting the most appropriate strategy to achieve highly efficient prognostics and decision-making.

A horizontal comparison indicates that no learning paradigm is uniformly superior, because its effectiveness depends on how well its underlying assumptions match the dominant characteristics of the available data. Incremental and online learning suit continuously evolving data streams because they emphasize knowledge updating and temporal adaptability, though they remain vulnerable to forgetting and error accumulation. Meta-learning prioritizes rapid few-shot adaptation by extracting transferable learning strategies from related tasks, whereas RL operates mainly at the decision level by optimizing sequential trade-offs among safety, maintenance cost, and operational availability. Graph-based and physics-informed architectures improve structural representation and physical consistency, but their performance depends on reliable sensor topology or prior physical knowledge. Therefore, algorithm selection should be determined by the dominant data limitation, temporal evolution pattern, and engineering objective rather than by isolated numerical performance indicators.

By integrating uncertainty-aware probabilistic outputs, targeted processing of ambiguous decision samples, and drift-resistant dynamic thresholding mechanisms, the adaptive decision-making framework can significantly reduce both false alarms and missed detections. Consequently, this comprehensive system robustly guarantees the engineering reliability of predictive maintenance systems operating within highly non-stationary environments.

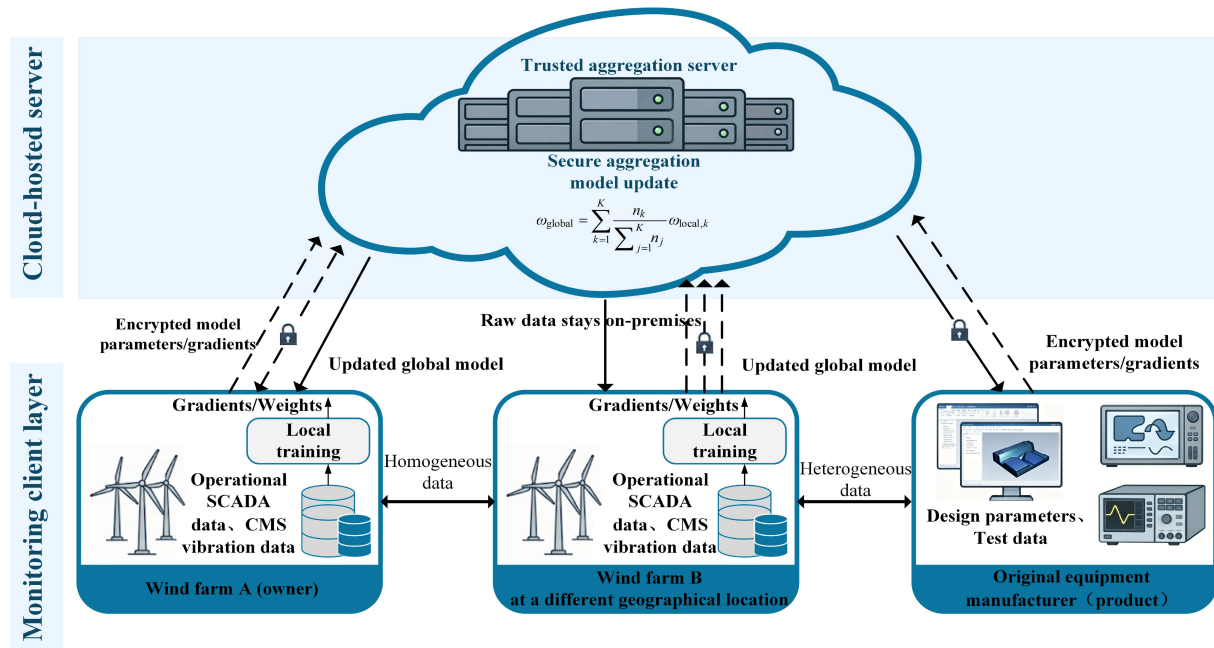


Figure 7. FL-based wind turbine O&M framework. FL: Federated learning; O&M: operation and maintenance; SCADA: supervisory control and data acquisition; CMS: condition monitoring system.

5. PRACTICAL CO-EVOLUTION AND TRUSTWORTHY O&M

This section examines the engineering deployment of intelligent O&M systems from four complementary perspectives: privacy-preserving federated collaboration, lightweight edge inference, trustworthy diagnostic decision-making, and digital-twin-oriented system integration. This work pays particular attention to trade-offs among privacy, computational efficiency, physical consistency, and lifecycle decision reliability.

5.1. Privacy-preserving collaborative mechanisms in FL

Although transfer learning provides algorithmic support for cross-farm generalization, privacy and commercial constraints often restrict direct sharing of raw fault data. FL addresses this limitation through a “data-stationary, model-mobile” mechanism: individual wind farms train models using local SCADA or vibration data, transmit encrypted parameters or gradients to a central server for aggregation, and receive the updated global model. The corresponding FL-based wind turbine O&M framework is illustrated in Figure 7. In Figure 7, “Raw data stays on-premises” means that the central aggregation server has no permission to access client raw datasets; only encrypted model parameters or gradients are exchanged. This process supports cross-farm knowledge transfer without exposing proprietary raw data, although communication efficiency, client heterogeneity, aggregation robustness, and adversarial security remain major deployment challenges.

Engineering practice has validated the value of collaborative innovation. Barber *et al.* showed that a wind-energy digital ecosystem can facilitate co-innovation and fault-detection collaboration; through the WinJi Gearbox Fault Detection Challenge and the EDP wind turbine fault detection challenge, participating solutions were associated with up to EUR 120,000 in O&M cost savings^[118]. This case illustrates the engineering value of distributed collaboration.

Fleet-level modeling and FL provide related but distinct approaches to collaborative diagnosis. A fleet-based anomaly-detection framework has been evaluated using wind-turbine vibration data^[119]. A separate fleet-wide model-generation process supports residual-based fault detection across wind turbines^[120]. These

Table 5. Challenges and countermeasures in FL for wind turbines

Core challenge type	Specific manifestations (wind power scenarios)	Representative strategies, algorithms
Data heterogeneity	Feature shifts across turbines, sites, and operating environments ^[122]	Clustered FL and client grouping
Data imbalance	Extreme fault scarcity and severe local class imbalance ^[123]	Prototype-based local balancing and weighted federated aggregation
Client heterogeneity and personalized adaptation	Differences among client data distributions motivate adaptive aggregation and client-specific models ^[124]	Dynamic weighting and dual-layer personalized FL

FL: Federated learning.

fleet-based approaches should not automatically be classified as FL. Explicit FL has also been investigated for privacy-preserving fleet-wide learning of wind-turbine conditions, allowing local model training without centralizing raw data^[121].

Although FL effectively dismantles physical data barriers, its practical deployment in the wind energy sector continues to encounter profoundly complex engineering challenges. Primarily, the pronounced disparities across diverse wind turbines-encompassing structural configurations, sensor setups, and operational environments-result in highly heterogeneous local data. This intrinsic heterogeneity severely undermines the convergence stability of the global model. Furthermore, the inherent long-tail distribution of wind turbine fault data is exacerbated within a distributed environment; consequently, minority fault classes are frequently inadequately represented at individual local nodes. To address these critical bottlenecks, the academic community has proposed a spectrum of targeted amelioration strategies. Table 5 systematically delineates the core challenges currently confronting FL in wind power applications, alongside their representative countermeasures.

To address these challenges, contemporary research is shifting from monolithic parameter averaging to personalized FL and meta-learning frameworks. By retaining localized, site-specific parameters atop globally shared knowledge, or by integrating cost-sensitive mechanisms to counteract local class imbalances, these systems can adapt to the specific operating conditions of individual wind farms while leveraging collective knowledge. In practice, the effectiveness of such frameworks depends on their ability to accommodate non-independent and identically distributed data, heterogeneous sensor configurations, and asynchronous communication across wind farms. Communication-efficient aggregation, privacy-preserving parameter exchange, and uncertainty-aware local adaptation are therefore essential for maintaining model stability under large-scale deployment. Moreover, dynamic client selection and continual model updating are required to prevent outdated local knowledge from degrading global diagnostic performance. Propelled by the convergence of hierarchical federated architectures and edge computing, future fault diagnosis and intelligent O&M systems are expected to evolve into a cloud-edge-physical-layer collaborative ecosystem for trustworthy full-lifecycle asset management.

5.2. Lightweight edge deployment

Edge deployment requires diagnostic models to satisfy strict constraints on latency, memory, energy consumption, and communication bandwidth. Accordingly, this section focuses on adaptive feature selection and model compression strategies that balance diagnostic accuracy with computational overhead under resource-constrained conditions^[125,126], while the broader cloud-edge-physical-layer architecture is discussed in Section 5.5.

5.2.1. Adaptive feature engineering at the edge

Processing high-frequency raw signals on resource-constrained edge devices can incur prohibitive latency. Edge feature engineering therefore aims to reduce computational and communication loads while retaining fault-sensitive information. Accordingly, the effectiveness of edge feature engineering should be evaluated not only by dimensionality reduction ratio, but also by its influence on diagnostic accuracy, inference latency, memory usage, and communication overhead.

Adaptive feature dimensionality reduction techniques are extensively deployed to alleviate severe communication pressure. Tailored for offshore wind power scenarios, recent studies propose executing primary feature extraction and selective screening directly at the edge gateway, thereby transmitting only the compact feature representations to the cloud. Edge-cloud collaboration can reduce raw-data transmission by moving preliminary processing closer to the data source, although the achieved bandwidth reduction depends on the task and deployment design^[127]. Complementary SCADA-based studies have combined fault detection with sensor or feature selection, showing that redundant variables can be reduced while retaining information relevant to turbine condition monitoring^[128].

Feature selection can be formulated as a multi-objective optimization problem involving a search for Pareto-efficient trade-offs between classification performance and feature-set size^[129,130]. For edge diagnosis, the selected solution must additionally be evaluated against latency and energy constraints. Moreover, the selected feature subset should remain sufficiently stable across varying wind speeds, loads, and operating regimes to avoid frequent feature reconfiguration during online deployment.

To this end, Han *et al.* proposed a multi-objective particle swarm optimization algorithm with adaptive strategies for feature selection^[131]. By jointly optimizing classification performance and feature-subset size, the method provides a relevant optimization basis for resource-constrained edge diagnosis. However, the effectiveness of edge-side feature engineering depends on reliable operating-state recognition and accurate estimation of device resource availability. Excessive dimensionality reduction may discard weak incipient-fault signatures, while frequent switching between feature subsets can add scheduling overhead and temporal inconsistency. Therefore, feature selection should be jointly optimized with lightweight model architectures to balance diagnostic accuracy, latency, communication cost, and energy consumption across the entire edge pipeline. Feature-level optimization alone, however, cannot fully satisfy the computational and memory constraints of embedded edge platforms, making model-level lightweight design and compression equally important.

5.2.2. Efficient architectures and model compression

Beyond data-level optimizations, the intrinsic lightweight design of deep learning models themselves is paramount for actualizing edge deployment. To rigorously evaluate the applicability of diverse models in resource-constrained environments, the academic community primarily focuses on the trade-off between model complexity and diagnostic performance.

As depicted in Figure 8, HRC-NASNet-b and WTBMobileNet occupy the low-complexity region and provide favorable accuracy-efficiency trade-offs. Their reported accuracies are lower than those of VGG-16 and MCWT-WCFormer in the figure; therefore, they should not be described as surpassing the heavyweight baselines. MCWT-WCFormer achieves the highest accuracy among the compared models but at substantially greater computational cost than the compact architectures, illustrating the performance-efficiency trade-off relevant to edge deployment. The technical pathways for lightweight wind turbine fault diagnosis can be broadly classified into two categories: (1) the design of inherently lightweight and

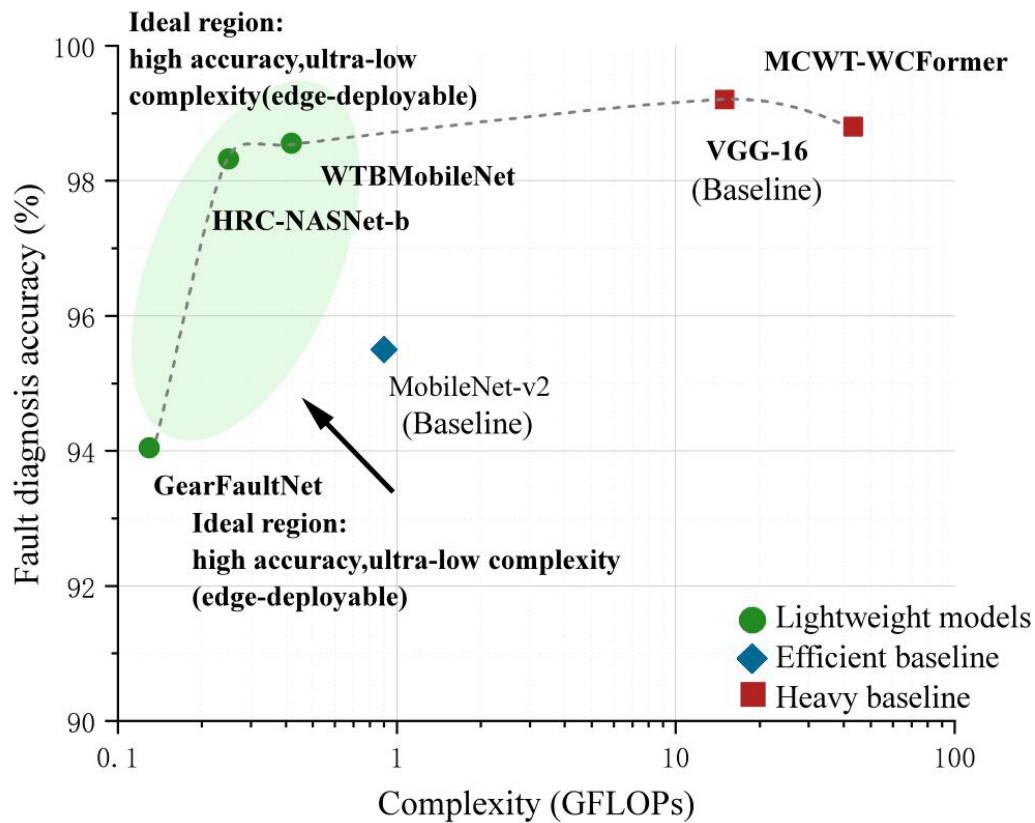


Figure 8. Accuracy–complexity trade-off of lightweight models and baselines. GFLOPs: Giga floating-point operations.

domain-specific network architectures; and (2) the compression and hardware-aware optimization of existing high-performance models through structured pruning, knowledge distillation, and neural architecture search (NAS).

These lightweight strategies improve deployment efficiency through fundamentally different mechanisms. Purpose-built lightweight networks reduce computational complexity through compact architectural design and therefore provide relatively predictable inference latency, but their limited representational capacity may limit recognition of weak or compound faults. Structured pruning exploits redundant channels, filters, or network blocks in over-parameterized models, although excessive pruning may remove fault-sensitive pathways and degrade diagnostic accuracy. Knowledge distillation transfers the soft decision relationships learned by a high-capacity teacher model to a compact student model and may preserve more diagnostic information than direct pruning, but its effectiveness depends strongly on teacher quality and teacher–student compatibility. Hardware-aware NAS explicitly incorporates latency, memory, and energy constraints into the optimization objective, enabling device-specific accuracy–efficiency trade-offs, although its high search cost and limited cross-platform portability constrain large-scale reuse.

For the first category, researchers have developed various domain-specific lightweight network architectures tailored to the distinct characteristics of wind turbine vibration signals. For instance, a customized lightweight one-dimensional CNN designed exclusively for real-time gearbox monitoring introduces asymmetric convolution kernels to reduce redundant computations. This architecture significantly reduces inference latency while maintaining a high recognition accuracy of 94.04%^[132].

Table 6. Quantitative comparison of lightweight fault diagnosis models

Model	Dataset	Reported performance	Model complexity
Lite CNN	CWRU	99.78%-99.98% accuracy	0.64% Params, 0.05% FLOPs, and 6.399% computation time relative to ResNet50 ^[137]
1D-2DIFCNN	CWRU; rotating-machinery fault-simulation test-rig data	100% accuracy; 96.29% average cross-condition accuracy	414.16 KB; 0.0486 s per training step ^[138]
LECA-EfficientNetV2	SEU gearbox dataset	99.38% bearing and 99.75% gear accuracy; 99.02%-99.63% transfer accuracy	Smallest parameter count among eight models; transfer diagnosis time of 9.58-9.92 s ^[139]

CNN: Convolutional neural network; FLOPs: floating-point operations.

Furthermore, in visual inspection, the WTBMobileNet architecture has been deployed for wind turbine blade surface defect recognition. Compared to the baseline network, this model achieves a 9.4-fold reduction in parameter count and a 2.7-fold decrease in computational overhead, accompanied by a marginal accuracy degradation of merely 1.68%. Such exemplary performance unequivocally demonstrates its exceptional engineering practicability^[133].

The second category focuses on compressing and optimizing existing high-performance yet over-parameterized models. Structured pruning systematically eliminates redundant convolutional channels, filters, or network blocks, thereby reducing computational overhead and facilitating efficient deployment on hardware acceleration platforms. For instance, in the Single Shot MultiBox Detector framework, replacing the VGG-16 backbone with ShuffleNetV2, coupled with a rigorous channel-pruning strategy, significantly increases the inference frame rate on embedded devices^[134].

Concurrently, knowledge distillation paradigms compel lightweight edge models to mimic the output distributions of cumbersome cloud-based models. This mechanism ensures compact networks inherit the robust generalization of complex models while maintaining a remarkably low parameter count, further optimizing the performance-efficiency trade-off shown in [Figure 8](#).

To further dismantle the barrier between algorithmic design and hardware deployment, cutting-edge research has actively pivoted toward hardware-aware NAS. Departing from the conventional paradigm that treats model design as an isolated algorithmic endeavor, this technology explicitly integrates the latency, energy consumption, and memory footprint of the target hardware as strict constraints within the neural network's search space^[135].

As a distinct architecture-design approach, MCWT-WCFormer combines wavelet-based multi-sensor processing with a CNN-Transformer network and reports approximately 4.2 ms for single-sample diagnosis under its experimental configuration^[136]. This result should not be attributed to NAS or treated as evidence of equivalent latency on an untested edge device. To complement the accuracy-complexity relationship illustrated in [Figure 8](#), [Table 6](#) provides a quantitative comparison of lightweight fault diagnosis models in terms of diagnostic performance and computational complexity.

These results indicate that lightweight performance should be evaluated jointly by accuracy, parameter count, computational operations, and execution time. Input compression and compact CNN design provide substantial efficiency gains, whereas feature-fusion and transfer-learning models generally introduce additional computation in exchange for improved robustness and cross-condition generalization.

Overall, lightweight edge deployment is achieved primarily through two complementary pathways: designing inherently efficient network architectures and compressing or hardware-aware optimizing existing

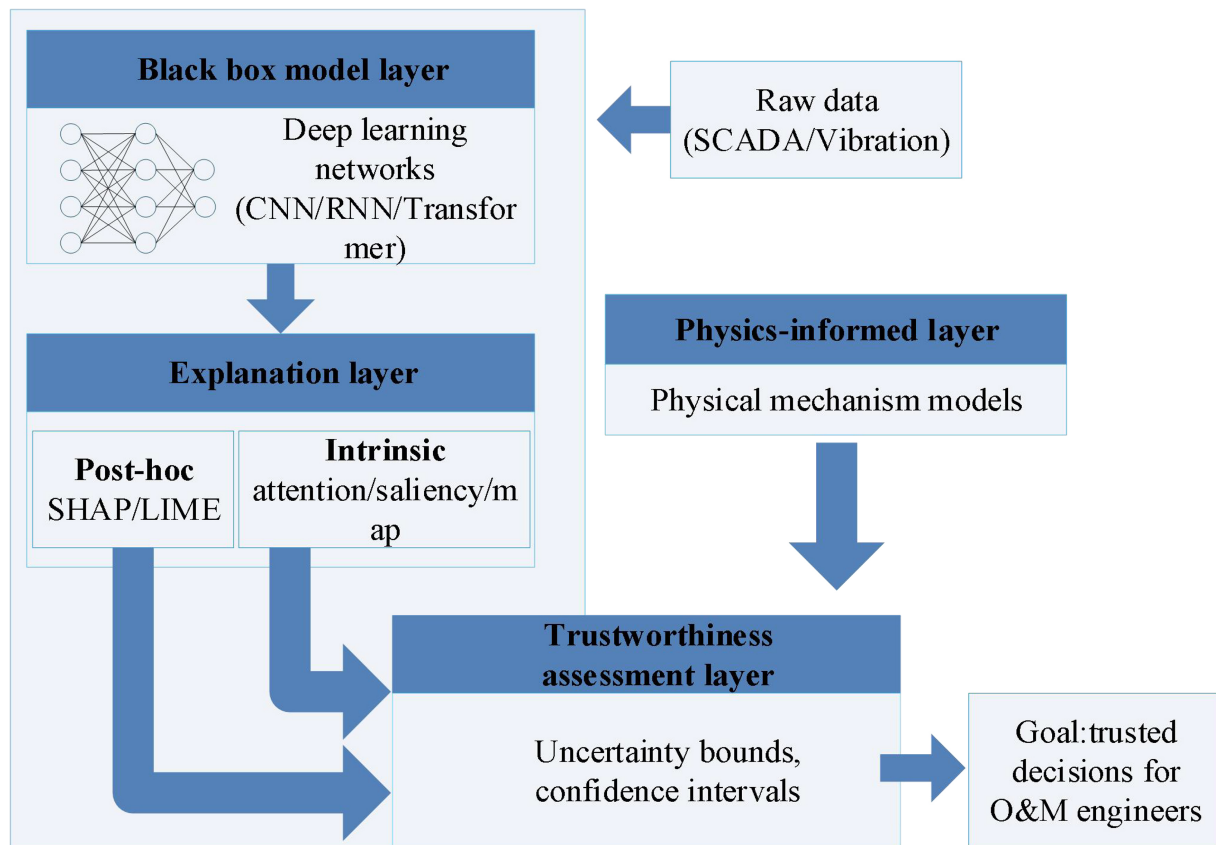


Figure 9. Framework for trustworthy wind turbine diagnosis. CNN: Convolutional neural network; RNN: recurrent neural network; SCADA: supervisory control and data acquisition; SHAP: Shapley additive explanations; LIME: local interpretable model-agnostic explanations; O&M: operation and maintenance.

high-performance models. In practice, selecting these strategies should consider diagnostic accuracy, inference latency, memory footprint, energy consumption, and compatibility with the target hardware platform^[140].

5.3. Trustworthy diagnosis: explainability and uncertainty quantification

Engineering deployment requires diagnostic models to provide not only accurate predictions but also transparent reasoning and quantified uncertainty. Because black-box alarms alone are insufficient to justify safety-critical actions such as turbine shutdowns, trustworthy diagnosis should integrate explainable AI, uncertainty quantification, and physical consistency, as illustrated in Figure 9.

5.3.1. Post-hoc interpretability in black-box diagnostics

To bridge the critical chasm between high-performance black-box models and actionable human decisions in wind power O&M, post-hoc explainability methods have emerged as indispensable tools. Their fundamental objective is to deconstruct the model's decision-making logic by rigorously quantifying feature contributions.

Within the complex, multi-variable coupled environment of wind turbine monitoring, model outputs typically depend on a vast array of interconnected operational parameters. Consequently, merely predicting an impending fault is insufficient to build trust among on-site engineers or formulate concrete intervention strategies^[141]. Methods such as Shapley additive explanations (SHAP) and the sensitivity analysis-based differential importance measure (DIM) address this challenge by assigning quantifiable impact scores to

individual input features. In doing so, they translate inherently opaque predictive outputs into auditable, evidence-based diagnostic rationales^[142].

SHAP analysis is utilized not only to identify novel, highly interpretable variable sets associated with rotor and pitch control for anomaly detection^[143], but also to dissect the deep correlations between inputs and hidden layer outputs within prognostic models. This facilitates localizing and eliminating low-contribution features, thereby optimizing sensor configuration and deployment strategies^[144]. Concurrently, the DIM quantifies the relative contribution of individual parameters or parameter groups to changes in a risk metric, thereby supporting sensitivity-based prioritization in risk-informed decision-making^[145].

The main advantage of these methodologies is twofold: they can precisely pinpoint critical fault indicators^[146] and, on the other hand, facilitate root cause localization and false alarm suppression by distinguishing anomalies caused by component degradation from deviations induced by external operating conditions.

Furthermore, such quantifiable interpretations can be directly integrated into decision support systems. By generating diagnostic rules based on a minimal set of relevant features, this approach provides acceptable explanations for anomalies and achieves detection accuracies up to 80%, thereby accelerating and standardizing O&M responses. Additionally, the underlying logical relationships unveiled by Shapley value analysis can be distilled into intuitive, comprehensible operational state indicators. This establishes a direct bridge between the model's inferential logic and empirical judgments made by site engineers on system behavior^[147].

Consequently, by systematically attributing model decisions to specific, measurable input variables, post-hoc explainability methods provide the requisite foundation of trust for the verification, comprehension, and execution of AI-driven diagnostics in high-risk maintenance scenarios.

5.3.2. Inherent interpretability and uncertainty estimation

Intrinsic interpretability depends on whether a model has an understandable structure and whether its explanations faithfully represent the prediction process. Attention mechanisms can expose feature-weighting patterns for inspection, but attention weights alone do not establish a transparent or causally valid diagnostic pathway.

For instance, a cosine-enhanced channel attention mechanism has been used to learn and accentuate effective features with explanatory value while suppressing noise interference^[148]. Meanwhile, parameter-free attention modules can integrate formalized prior knowledge of outlier types into CNNs. This achieves a seamless fusion of domain knowledge and data-driven learning, supporting feature selection and diagnostic interpretation^[149]. Pushing architectural innovation further, recent approaches constrain standard dot-product attention using learnable adaptive Gaussian mixture models. By explicitly targeting the extraction of local sparse features and mitigating sensitivity to irrelevant information, these methods provide additional structure for inspecting feature extraction, without by themselves guaranteeing faithful explanations^[150].

Beyond architectural innovations, feature visualization techniques translate internal model representations into human-comprehensible formats, serving as a visual bridge between complex data manifolds and engineering intuition.

A prominent approach employs a supervised VAE to project high-dimensional wind turbine states into a lower-dimensional latent space. This generates 2D trajectory maps that visualize health degradation processes and facilitate early anomaly detection, with studies demonstrating the capability to identify main

bearing over-temperature signatures up to 11 days prior to control system alarms^[151]. Similarly, dimensionality reduction and visualization techniques, such as t-distributed stochastic neighbor embedding, have been integrated into advanced diagnostic frameworks like the long-term feature memory network and deformable spatio-temporal fusion network (DSTF-Net). These tools intuitively illustrate the distribution of fault features learned from complex insulated bearing data. These visualizations can illustrate separation within the learned feature space, but do not establish physical validity, causal interpretation, or calibrated confidence. Corresponding studies report accuracy enhancements ranging from 4.2% to 29.7%^[152,153].

Attention-assisted inspection and feature visualization can help analysts examine learned representations. Their explanatory value should nevertheless be assessed separately from diagnostic accuracy, using appropriate checks of explanation fidelity, stability, and physical plausibility.

5.3.3. Bayesian and probabilistic deep learning frameworks

Beyond achieving decision explainability, establishing genuine engineering trust necessitates a transition from qualitative comprehension to a quantitative foundation—specifically, the explicit quantification of predictive uncertainty inherent in data-driven models to transform heuristic forecasts into calibrated risk metrics. Central to this pivotal transition is the deployment of probabilistic and Bayesian deep learning frameworks. Rather than yielding singular point estimates, these architectures output a probability distribution of potential outcomes, thereby formally articulating the confidence of the model in its internal inferences.

For instance, uncertainty-aware Bayesian optimization applied to extreme recurrent networks explicitly integrates predictive uncertainty into the objective function to be minimized. This approach yields superior generalization compared to deterministic counterparts, with the coefficient of determination improving by $5.14\% \pm 2.13\%$ ^[154]. Furthermore, the synergy of parallel gated recurrent units, kernel density estimation, and Monte Carlo dropout facilitates the generation of probabilistic RUL estimates. This ensemble preserves high diagnostic precision while simultaneously delineating reliable uncertainty bounds^[155]. The selection of probabilistic models also profoundly impacts performance; in early-stage fault detection, Gaussian mixture copula models have demonstrated superiority over alternative probabilistic power curve paradigms, accelerating detection speed by a factor of 10-30^[156].

Jing and Zhao developed an adjustable piecewise multivariable regression strategy that groups operating-data bins with similar nonlinear characteristics and fits Gaussian process regression submodels within the resulting clusters. The associated probabilistic monitoring strategy uses the predictive distribution to quantify condition uncertainty and reports a wind-turbine abnormal-condition detection accuracy of 95.35%^[157].

Furthermore, incorporating temporal dependencies and prior knowledge can substantially enrich such point-wise uncertainty estimation. For instance, a Bayesian recurrent state estimator, when integrated with historical degradation trends for robust RUL prediction, can reduce prediction errors by over 16%^[158]. Similarly, variational deep Gaussian processes strengthen uncertainty representation within multi-layer architectures while concurrently identifying low-contribution features to optimize sensor deployment strategies.

Providing robust theoretical underpinning, Bayesian networks and related probabilistic graphical models have been systematically reviewed as powerful tools for characterizing uncertainty and complex dependencies, directly addressing both operational and failure uncertainties within wind energy systems^[159]. Advancing this paradigm, sophisticated hybrid models, such as D-vine copula Bayesian networks, model

fault correlations across multi-channel condition monitoring data. This enables the rigorous quantification of component risk probabilities and maintenance priorities, successfully overcoming the simplistic assumption of mutually independent inputs^[160].

Ultimately, upon completing the uncertainty modeling via Bayesian or probabilistic deep learning, the practical viability of the output of the model is inherently contingent upon whether the confidence levels can be adequately calibrated, and whether the decision-making layer can leverage this probabilistic information to execute risk-aware maintenance scheduling.

Ultimately, whether by embedding Gaussian processes into network output layers to formulate health-aware distributions^[161], leveraging non-Gaussian deep state-space models equipped with planar flows to relax rigid distributional assumptions^[162], or adopting the broader spectrum of probabilistic methodologies prevalent in statistical fault detection—such as probabilistic principal component analysis and probabilistic partial least squares—uncertainty quantification unequivocally establishes itself as the core quantitative pillar^[163]. It underpins trustworthy, risk-aware decision-making in the field, thereby propelling diagnostic systems beyond mere interpretability toward intrinsic credibility.

5.3.4. Confidence calibration and decision fusion

Building on uncertainty quantification, the engineering reliability of predictive outputs hinges not merely on generating probabilistic estimates, but on whether these probabilities are well calibrated and translated into actionable decisions through strategic fusion. This is paramount for managing missed detections and false alarms in dynamic environments. In practical windfarm deployment, miscalibrated model confidence often leads to two harmful practical outcomes: overconfident alarms for normal operational fluctuations and missed alerts for incipient, low-amplitude component degradation. Furthermore, the inherent volatility of wind turbine operations, particularly across diverse operating conditions, often induces discrepancies between model confidence and actual accuracy, thereby necessitating specialized calibration techniques for realignment.

For instance, to address fluctuating reconstruction errors under varying operating conditions, Xiao *et al.* proposed the SME index for multi-condition fault detection, making anomaly evaluation more consistent and comparable than evaluation based on raw reconstruction loss^[164]. Furthermore, to prevent regression models—such as power curves—from being contaminated by anomalous data, Zhang *et al.* introduced a risk-regularized optimization method formulated as a chance-constrained problem. This approach effectively filters outliers while retaining normal data with high probability, eliminating the need for external reference datasets or extensive hyperparameter tuning^[165].

Beyond single-model calibration, uncertainty-aware models can further elevate diagnostic credibility. Amin *et al.* developed a Bayesian CNN for wind turbine gearbox diagnosis that flags unfamiliar inputs through elevated predictive uncertainty, reducing the risk of overconfident misclassification^[166]. Complementarily, Pandit and Infield constructed Gaussian-process operational curves that provide both mean estimates and 95% confidence intervals for wind turbine condition monitoring^[167]. For early fault diagnosis, Zhao proposed a threshold-free strategy that leverages the probability density function of estimation errors alongside self-organizing neural networks to cluster fault features, yielding a diagnostic accuracy exceeding 90% for early pitch system faults^[168]. Taken together, these approaches indicate that reliable wind-turbine diagnosis requires not only accurate predictions, but also calibrated confidence estimates and uncertainty-aware decision thresholds.

Table 7. Comparative analysis of trustworthy intelligent diagnosis technologies for wind turbines

Diagnostic paradigm	Representative techniques	Core advantages	Inherent limitations	Typical wind O&M scenarios
Post-hoc interpretability	SHAP feature attribution	Quantifies feature contributions to damage indices and helps interpret operational influences ^[169]	Explains model outputs; does not establish physical causality	Interpreting damage indices for wind-turbine blade damage detection
Attention-assisted feature inspection	Channel–space attention; feature visualization	Enhances multiscale feature discrimination for gearbox compound-fault diagnosis ^[170]	Attention weights and visual clusters do not guarantee faithful or causal explanations	Inspecting learned gearbox fault representations; wind-field validation is still required
Uncertainty quantification	Temporal convolutional variational deep Gaussian processes	Produces RUL confidence intervals and explicitly quantifies predictive uncertainty ^[171]	Probabilistic inference increases computational complexity	Wind-turbine gearbox RUL prediction under variable operating conditions
Decision thresholding	Counter-based residual thresholding	Balances missed detections against false alarms and improves sensitivity to small faults ^[172]	Threshold and counter settings require scenario-specific tuning	Sensor, actuator, and drivetrain fault detection
Physics-informed modeling	PINNs	Incorporates governing equations into learning for forward and inverse PDE problems ^[173]	Requires valid governing equations and careful loss balancing	Methodological foundation for potential wind O&M applications; not wind-specific validation

O&M: Operation and maintenance; SHAP: Shapley additive explanations; RUL: remaining useful life; PINNs: physics-informed neural networks; PDE: partial differential equation.

To further delineate the application boundaries and performance trade-offs of the aforementioned trustworthy diagnostic technologies in engineering practice, [Table 7](#) provides a systematic taxonomy. It categorizes these approaches across key dimensions: representative techniques, core advantages, inherent limitations, and typical application scenarios within the wind power sector. A comparative analysis reveals a pronounced trade-off between the depth of interpretation and computational overhead across different methodologies. Consequently, this necessitates that O&M decision-making systems flexibly select or fuse appropriate technological pathways contingent upon specific monitoring targets and real-time requirements.

In summary, through error metric calibration, risk-regularized learning, and UQ-driven decision fusion, these methodologies translate raw probabilistic outputs into robust, actionable O&M insights, thereby closing the critical loop between predictive uncertainty and reliable maintenance actions.

5.3.5. Integrating physics-based models with counterfactual analysis

The ultimate objective of trustworthy condition monitoring extends beyond merely identifying statistical correlations; it lies in unraveling the underlying causal mechanisms of faults. This is achieved by coupling data-driven pattern discovery with physical laws to provide interpretations with profound intrinsic insights. Although advanced data-driven models excel at detecting anomalies and associating them with specific signal patterns, their diagnostics frequently remain at the correlational level, making them susceptible to confounding by variations in external operating conditions. To overcome the limitation that statistical associations in purely data-driven methods lack physical causal constraints, a vital research trajectory involves incorporating explicit physical models into the diagnostic framework. For instance, addressing the ambiguity of vibration signatures in planetary gearboxes, research has formulated a novel transfer path effect model based on modulation phenomena. This physical model can generate interpretable phenomenological signals for both healthy and faulty states, thereby furnishing a priori guidance and causal-mechanistic explanatory bases for monitoring decisions, transcending the diagnostic logic of pure data correlation^[174].

In parallel, a complementary class of data-driven strategies attempts to distinguish influential associations within learned models by analyzing sensitivity to input perturbations. Model-agnostic explanation methods can quantify how changes in input variables influence a learned prediction^[175]. Such perturbation-based explanations can identify influential variables, but they should not be interpreted as causal effects without additional causal assumptions or interventions^[176].

Physics-based interpretation requires substantial domain knowledge and may not encompass unmodeled dynamics. In contrast, perturbation-based explanations remain dependent on the accuracy and completeness of the underlying detection model and do not provide the mechanistic depth of first-principles models. Overall, these two pathways—one embedding physical causality into the model and the other extracting explanatory signals from model behavior—are complementary. Together, they support more credible O&M decisions while clarifying the limits of causal interpretation.

5.4. Future-oriented full-lifecycle intelligent O&M decision frameworks

The ultimate value of intelligent fault diagnosis systems extends far beyond precisely identifying faults; it lies in leveraging trustworthy diagnostic insights to assist in formulating optimal O&M strategies, thereby minimizing the full-lifecycle LCOE. With the continuous expansion of wind farm scales and the increasing service life of turbine units, O&M paradigms are transitioning from singular reactive or scheduled maintenance toward condition-based predictive maintenance, and ultimately, prescriptive maintenance. This paradigm shift dictates that intelligent systems must shatter the limitations of focusing solely on isolated faults, necessitating the construction of comprehensive decision-making architectures that encompass the equipment's full lifecycle and multi-component coupling dynamics.

5.4.1. Integrated management spanning the full lifecycle and multi-component systems

A well-established standardized O&M framework must transcend reactive response modes, instituting a full-lifecycle management paradigm—from design through decommissioning—to enhance system sustainability and efficiency^[177]. This macroscopic philosophy must be concretely implemented at the microscopic component level. For instance, addressing bearing lubrication failures requires constructing a comprehensive scheme that integrates failure analysis, tribological mechanisms, and intelligent lubrication^[178-180]. For generators, AI technologies have already permeated the entire process spanning their design and maintenance^[181]. Meanwhile, for offshore wind support structures, harsh operating conditions impose more stringent demands on monitoring robustness^[182]. However, this architecture, fusing temporal dimensions with component depth, significantly amplifies system complexity, necessitating the capability to seamlessly integrate heterogeneous domain knowledge and multi-source data. Ultimately, only by achieving a deep synergy between full-lifecycle tracking and critical component management can O&M frameworks truly fulfill the dual demands for reliability and economic viability in modern wind power systems.

5.4.2. Construction and advanced optimization of O&M decision support systems

The core of a standardized O&M framework lies in its decision support system. This system is no longer confined to simplistic condition alarms; rather, it aims to integrate front-end equipment health prognostics, mid-end maintenance strategy optimization, and back-end resource scheduling into a unified decision engine. The paramount challenge for the decision support system (DSS) resides in outputting O&M strategies that are optimal in terms of both cost and risk, amidst a multitude of compounding uncertainties—including model prediction errors, economic parameter fluctuations, logistical constraints, and the inherent stochasticity of component degradation. Architecturally, a DSS typically encompasses two tightly coupled modules: health prognostics and maintenance decision-making^[183]. Its transition from theoretical constructs to practical engineering applications fundamentally represents an algorithmic evolution from deterministic planning toward explicit uncertainty modeling.

Table 8. Comparative analysis of optimization algorithms for O&M decision support

Application scenario	Algorithmic methodology	Key addressed O&M pain points
Multi-objective maintenance planning	NSGA-II with heuristic maintenance-team scheduling and an (s, S) inventory policy	Balancing annual production losses and maintenance costs under wind-speed, crew, and spare-parts constraints ^[184]
Condition-based maintenance strategy	Distributionally robust chance constraints	Mitigating decision failures from RUL prediction errors and uncertainties ^[185]
Multi-asset dynamic scheduling	Deep RL ensemble using domain-informed DQNs for offshore wind farm maintenance scheduling	Coordinating maintenance scheduling decisions and turbine selection under wake effects and variable weather conditions to improve operational efficiency ^[186]
Single-turbine short-term predictive maintenance	GPR-based probabilistic RUL prediction, and adaptive differential evolution (JADE) for maintenance optimization	Differential evolution-based optimization for single-turbine short-term predictive maintenance under probabilistic RUL scenarios, considering electricity and component prices, wind speeds, and maintenance duration ^[187]
Life extension of aging turbines	Multi-criteria lifetime-extension assessment	Integrating structural condition, safety, economic viability, and regulatory requirements when deciding continued operation or decommissioning ^[188]

O&M: Operation and maintenance; NSGA-II: non-dominated sorting genetic algorithm II; RUL: remaining useful life; RL: reinforcement learning; DQNs: deep Q-network; GPR: Gaussian process regression; JADE: adaptive differential evolution.

Different O&M optimization methods exhibit distinct advantages because they represent uncertainty, conflicting objectives, and sequential decisions in fundamentally different ways. Evolutionary multi-objective optimization can balance production losses and maintenance costs, robust optimization emphasizes protection against uncertain degradation predictions, and reinforcement-learning-based methods are more appropriate for high-dimensional dynamic scheduling but depend strongly on simulator fidelity and reward design. Table 8 systematically summarizes the key algorithms embedded within the DSS and the specific O&M pain points they target.

The engineering value of integrating these algorithms has been corroborated by multiple empirical studies. Simulation results indicate that advanced DSS frameworks incorporating semi-Markov models and hierarchical failure analysis offer substantial economic advantages over traditional strategies^[189,190]. For example, a reference model fusing preventive and predictive maintenance demonstrates a potential reduction in overall maintenance costs by up to 32%^[191]. Concurrently, quantitative evaluation methods tailored for short-term cost-effectiveness further validate the economic rationality of such decision outputs. Ultimately, by standardizing the encapsulation of advanced algorithms—such as evolutionary optimization, robust optimization, and deep RL—within the decision kernel, modern DSS effectively transform the inherent uncertainties of wind farm operations into quantifiable and controllable economic risks. This constitutes the critical nexus bridging data and value within the intelligent O&M architecture.

5.5. From architectural synergy to algorithmic enhancement

As wind turbines evolve toward larger capacities and deep-offshore environments, traditional monitoring paradigms struggle to keep pace with complex system evolution. Digital twin technology provides a new paradigm for full-lifecycle management by constructing dynamic mappings of physical entities. This section analyzes the latest advancements across three dimensions: cloud-edge-physical-layer architectural synergy, spatio-temporal modeling algorithms, and underlying physical data support.

5.5.1. Cloud-edge-physical-layer collaborative architecture and IoT communication foundations

Realizing fleet-level collaborative monitoring fundamentally relies on constructing a highly reliable Internet of Things (IoT) framework and a flexible computational scheduling system. The paradigm shift from single-turbine monitoring to networked synergy depends heavily on a communication infrastructure capable of supporting the seamless convergence of heterogeneous data streams. For instance, in a real-world

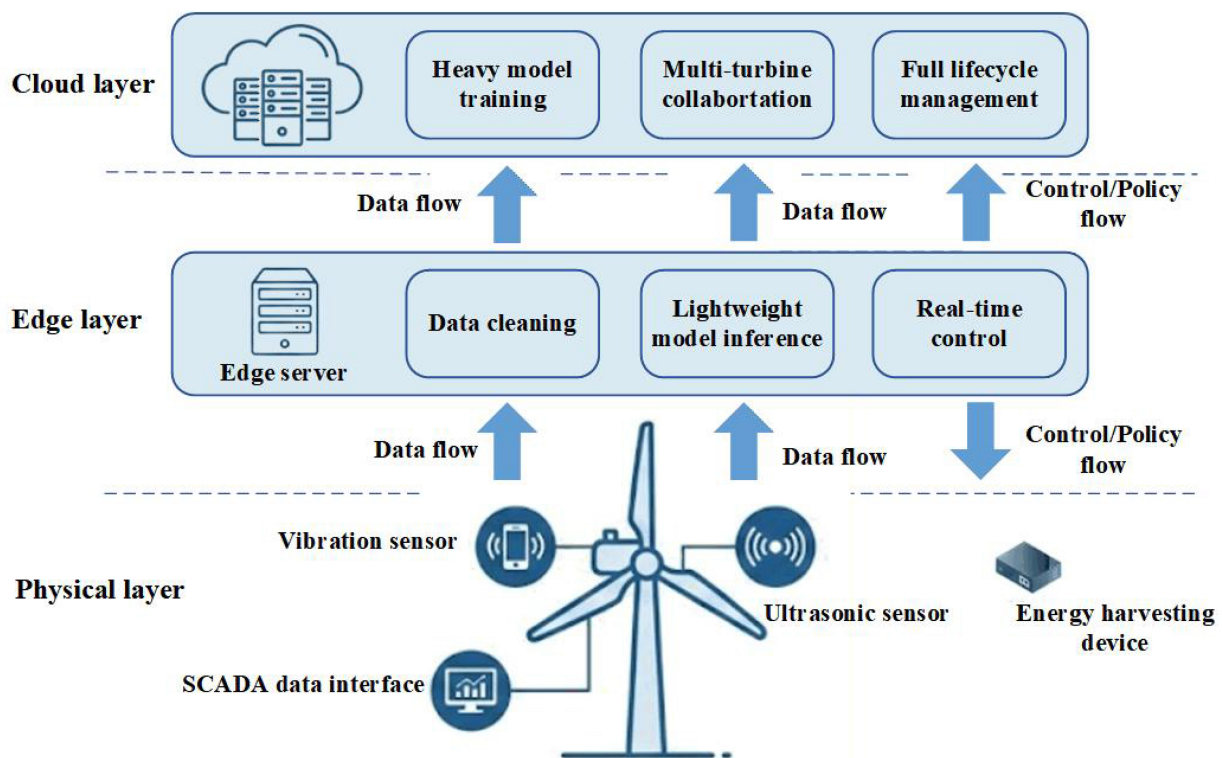


Figure 10. Cloud-edge-physical-layer collaborative O&M architecture for wind turbines. O&M: Operation and maintenance; SCADA: supervisory control and data acquisition.

deployment involving 75 urban turbines, an IoT network comprising 2,300 sensors maintained a high mean time between failures while facilitating deep learning-based wind condition forecasting and system optimization. This implementation yielded substantial benefits, notably a 41% reduction in maintenance costs and a 23.8% increase in annual energy production^[192]. Figure 10 illustrates a typical tri-tier cloud-edge-physical-layer smart O&M architecture along with its internal data flow dynamics. Through hierarchical decoupling, this architecture achieves a structured integration of massive data streams derived from physical sensing with the control streams originating from cloud-based decision-making.

As illustrated, the core advantage of this architecture lies in the differentiated deployment of functions and closed-loop synergy: the foundational layer focuses on the holographic sensing of multi-source heterogeneous data; acting as the first line of defense for real-time operations, the edge layer undertakes data cleaning and lightweight inference tasks to alleviate uplink bandwidth pressure; meanwhile, the cloud leverages its computational superiority to handle heavy-duty model training and dispatches the generated strategies down to the edge. Based on this hierarchical division of labor, the cloud-edge-physical-layer collaborative architecture resolves edge resource constraints through dynamic task offloading. For offshore wind scenarios, a task-offloading strategy based on multi-agent deep RL can reduce system overhead by approximately 61.8%^[193]. At the same time, this architecture enables a shift from anomaly detection to comprehensive decision support. For instance, a fuzzy inference system utilizing multispectral imagery can output maintenance priority scores highly correlated with expert assessments^[194]. Furthermore, for time-sensitive tasks such as low-voltage ride-through, neural predictive control at the edge achieves millisecond-level responses^[195]. Here, millisecond-level response refers to control-loop inference and actuation latency for time-critical electrical control tasks, rather than the prognostic lead time of mechanical fault warnings, which is generally measured over substantially longer horizons. Theoretical research also underscores that

situational awareness, integrating state-space models and AI, is crucial for realizing system-level adaptive decision-making^[196].

5.5.2. Multi-scale spatio-temporal modeling and adaptive diagnostic algorithms

The high fidelity of digital twins relies heavily on algorithms striking an optimal balance between real-time responsiveness and accuracy. For abrupt electrical faults, short-time signal amplitude ripple analysis and the three-stage temporal window mechanism have proven highly effective in extracting fault features^[197,198]. Furthermore, when addressing complex spatio-temporal coupling characteristics, deep learning models demonstrate outstanding performance: a residual LSTM integrated with an attention mechanism achieves exceptionally high precision in critical variable prediction^[199], while a self-attention LSTM attains a 98.67% accuracy rate in yaw system diagnostics^[200].

To address non-stationary operational data, spatio-temporal graph neural networks and interactive spatio-temporal networks enhance anomaly recognition capabilities through topological correlations. Meanwhile, wavelet LSTM networks and the differentiated frequency-spatial-temporal feature extraction with differential entropy method effectively mitigate the interference of signal non-stationarity on early detection. Furthermore, integrating the exponentially weighted moving average and adjacent-state adaptive threshold mechanisms can significantly reduce false alarm rates^[201-205], ensuring the robustness of decision-making. However, regardless of the algorithms' sophistication, their practical deployment efficacy at the edge is ultimately constrained by input data quality and physical sensing stability^[206-208]. This compels us to shift our focus downward from the algorithmic layer to the physical layer, re-examining sensing-system maintenance and data governance as the foundation of reliable intelligent O&M.

5.5.3. Physical sensing maintenance and data governance foundation

The sustainable operation of edge intelligence hinges on the dual guarantees of physical power supply and data quality. Addressing the energy supply bottlenecks of hard-to-maintain components, electromagnetic and pendulum-based energy harvesting technologies have emerged as critical breakthroughs. Specifically, electromagnetic harvesters inside the blades can output 6 mW of power to drive vibration monitoring^[209], whereas pendulum structures mounted on rotating shafts can achieve an output of 104.5 mW^[210], laying the physical foundation for self-powered wireless sensor networks. At the data level, semi-supervised multivariate models can exploit abundant unlabeled SCADA records together with limited labeled anomalies to improve wind-turbine anomaly detection^[211]. Meanwhile, sequence-driven automatic label calibration methods effectively resolve mislabeling issues^[212]. Simultaneously, to accommodate constrained edge computing power, Autoregressive compression techniques reduce dimensionality while heightening sensitivity to fault features^[213]. Coupled with covariate adjustment strategies^[214], these methods ensure that the data streams fed into the models possess high integrity and adaptability to varying operational conditions. Ultimately, these preprocessing technologies constitute the indispensable data bedrock for the digital twin ecosystem.

6. CONCLUSIONS AND FUTURE PERSPECTIVES

Centering on the prevalent data dilemmas in condition monitoring and fault diagnosis of wind power equipment—namely, the critical challenges of scarce high-value fault samples, extreme class imbalance, and uneven quality of multi-source heterogeneous data—this paper systematically reviews the research progress of data-driven intelligent O&M technologies by combining bibliometric analysis with state-of-the-art reviews. Throughout the paper, it is evident that current research has progressively evolved from local diagnostic strategies relying on single models or single data sources into multi-dimensional solutions encompassing data augmentation, algorithmic optimization, and system synergy. At the data level, generative methods represented by GANs, VAEs, and diffusion models have played a vital role in filling feature space voids and

mitigating fault sample scarcity, effectively improving the training stability and classification performance of deep models to a certain extent. However, these methods still encounter bottlenecks in balancing high fidelity with diversity, and their reliance on majority-class distributions restricts their complete credibility in extremely small-sample scenarios.

Meanwhile, algorithmic and feature-level innovations have forged novel pathways to tackle data scarcity. Semi-supervised and contrastive learning paradigms, by thoroughly mining the intrinsic topological structures of unlabeled data, have effectively reduced reliance on manually annotated samples, demonstrating excellent feature extraction capabilities in tasks like early fault diagnosis for gearboxes and bearings. Conversely, cost-sensitive learning, resampling techniques, and ensemble learning strategies offer effective means for classifier design with imbalanced data by directly rectifying the class bias of decision boundaries. Coupled with multi-domain feature enhancement and signal processing technologies, these form the cornerstone for improving model robustness under complex, non-stationary conditions. Furthermore, addressing the industry-wide issues of data silos and lagging model updates, federated and incremental learning are reconstructing the smart O&M ecosystem; the former enables cross-site knowledge sharing while safeguarding privacy, whereas the latter preliminarily resolves the catastrophic forgetting dilemma during full-lifecycle service.

Looking ahead, the field must move beyond isolated model improvements toward physics-informed, continuously adaptive, cross-domain, trustworthy, and deployment-oriented intelligence. The following six directions summarize the most actionable priorities for next-generation wind turbine O&M.

To further clarify the actionable pathways toward next-generation intelligent O&M systems, future research efforts can be summarized in the following specific directions:

(1) Physics-informed and interpretable intelligent diagnosis

Future research should focus on developing physics-data hybrid intelligence by integrating turbine dynamic models, degradation mechanisms, and expert knowledge into deep learning architectures. Such approaches are expected to improve the interpretability of diagnostic results and enhance model reliability under unseen fault conditions.

(2) Self-evolving learning paradigms under dynamic operating environments

Considering the long-term degradation characteristics and continuously changing operating conditions of wind turbines, future intelligent O&M systems should evolve from static models toward self-adaptive learning frameworks. Future studies should investigate continual learning algorithms capable of updating knowledge without catastrophic forgetting, automatic concept drift detection mechanisms for identifying distribution changes caused by environmental variations and component aging, and adaptive model updating strategies that balance learning capability, stability, and computational efficiency.

(3) Generalizable cross-domain intelligence for diverse wind farms

Although transfer learning and domain adaptation have significantly improved diagnostic model adaptability, large-scale deployment across different wind farms remains challenging because turbine structures, geographical environments, and operating conditions vary widely. Future research should focus on developing robust cross-domain intelligence approaches, including universal domain adaptation methods with limited target-domain information and causal representation learning techniques to separate intrinsic fault characteristics from environmental disturbances.

(4) Multimodal foundation models and large-scale knowledge utilization

As heterogeneous monitoring data accumulate, developing industrial foundation models for wind power is a promising research direction. A primary engineering challenge lies in designing modality-aware tokenization strategies for high-frequency vibration data streams. Rather than directly treating individual sampling points as tokens, future models should investigate multi-resolution temporal patching, time-frequency patch encoding, or vector-quantized event tokens to preserve transient impacts, phase information, and fault-sensitive spectral structures while controlling sequence length and facilitating alignment with low-frequency SCADA variables. For SCADA time-series data, self-supervised pretraining can be formulated as a masked state prediction task, in which selected sensor channels, temporal segments, or operating-condition states are masked and reconstructed from the remaining contextual information, enabling the model to learn dependencies among operating conditions, component degradation, and fault evolution from large-scale unlabeled data. By combining compact vibration tokenization with SCADA-based masked state prediction, foundation models can be pretrained across large turbine fleets and subsequently adapted to different turbine types and wind farms through parameter-efficient fine-tuning, thereby reducing annotation requirements, computational costs, and barriers to large-scale deployment.

(5) Cloud-edge-physical-layer collaborative intelligent O&M systems

To satisfy the requirements of real-time monitoring and large-scale industrial deployment, future research should investigate cloud-edge-physical-layer collaborative architectures. Lightweight diagnostic models should be developed for edge devices to achieve rapid anomaly detection under limited computational resources, while cloud-based models should undertake complex trend prediction, knowledge updating, and global optimization. Furthermore, the integration of digital twins with cloud-edge intelligence should be explored to establish closed-loop systems covering condition monitoring, fault diagnosis, RUL prediction, and maintenance optimization.

(6) Trustworthy and autonomous decision-making frameworks

Future intelligent O&M systems should extend beyond fault identification toward autonomous decision-making and maintenance optimization. Key research directions include combining RL with digital twins to optimize maintenance scheduling under uncertainty, developing risk-aware decision models that consider safety requirements, economic costs, and RUL, and establishing comprehensive evaluation frameworks to assess model reliability, interpretability, robustness, and industrial applicability before large-scale deployment.

In summary, future intelligent O&M research should shift from isolated algorithm improvement toward lifecycle-oriented, trustworthy, and adaptive intelligence. The ultimate goal is to develop autonomous wind power systems capable of continuous learning, reasoning, and reliable decision-making under complex and uncertain operating environments.

DECLARATIONS

Authors' contributions

Conceived and designed the review: Li, X.; Kong, Y.; Wang, T.

Conducted the literature searches, organized the references in Zotero, and compiled the bibliographic data: Li, X.; Wang, Z.

Performed the bibliometric analysis and prepared the figures and tables: Li, X.; Wang, Z.; Hu, W.

Analyzed and interpreted the reviewed literature: Li, X.; Hu, W.; Kong, Y.; Wang, T.; Qin, Z.; Chu, F.

Drafted the manuscript: Wang, Z.

Critically reviewed and revised the manuscript for important intellectual content: Wang, Z.; Hu, W.; Kong, Y.; Wang, T.; Qin, Z.; Chu, F.

Supervised the research: Kong, Y.; Wang, T.; Chu, F.

All authors read and approved the final manuscript.

Availability of data and materials

Not applicable.

AI and AI-assisted tools statement

Not applicable.

Financial support and sponsorship

This work is supported in part by the National Natural Science Foundation of China (52505091 and 52575094).

Conflicts of interest

Wang, T., Li, X. and Hu, W. are the Guest Editors of the Special Issue entitled “Artificial Intelligence-Based Fault Diagnosis and Intelligent Operation and Maintenance of Critical Equipment Components” in *Intelligence & Robotics*. Wang, T., Li, X. and Hu, W. were not involved in any stage of the editorial process, including reviewer selection, manuscript handling, and decision-making. The other authors declare no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2026.

REFERENCES

1. Ayman, A.; Onsy, A.; Attallah, O.; Brooks, H.; Morsi, I. Feature learning for bearing prognostics: a comprehensive review of machine/deep learning methods, challenges, and opportunities. *Measurement* **2025**, *245*, 116589. DOI
2. Elshenawy, L. M.; Gafar, A. A.; Awad, H. A.; Abouomar, M. S. Fault detection of wind turbine system based on data-driven methods: a comparative study. *Neural. Comput. Appl.* **2024**, *36*, 10279-96. DOI
3. Tang, G.; Cheng, L.; Xu, Z.; et al. Multimodal data enhanced wind generator fault diagnosis considering small sample conditions. *IEEE. Trans. Ind. Electron.* **2025**, *72*, 14831-42. DOI
4. Tang, M.; Tang, J.; Wu, H.; et al. Abnormal detection of wind turbine converter based on CWGANP-CSSVM. *Fractals* **2023**, *31*, 2340139. DOI
5. Xiangjun, Z.; Lingqin, X.; Chen, F.; Ming, Y. An improved conditional generative adversarial method for wind turbine gearbox fault diagnosis with imbalance data. *IEEE. Sensors. J.* **2024**, *24*, 38727-39. DOI
6. Han, T.; Xie, W.; Pei, Z. Semi-supervised adversarial discriminative learning approach for intelligent fault diagnosis of wind turbine. *Inf. Sci.* **2023**, *648*, 119496. DOI
7. Wang, Z.; Qin, B.; Sun, H.; et al. An imbalanced semi-supervised wind turbine blade icing detection method based on contrastive learning. *Renew. Energy.* **2023**, *212*, 251-62. DOI
8. Liu, J.; Li, X.; Li, C.; et al. When deep normal behavior models meet fault samples: a generalized wind turbine anomaly detection scheme. *IEEE. Trans. Instrum. Meas.* **2023**, *72*, 1-16. DOI
9. Wang, L.; Dai, S.; Kang, Z.; Han, S.; Zhang, G.; Liu, Y. A CNN-SA-GRU model with focal loss for fault diagnosis of wind turbine gearboxes. *Energies* **2025**, *18*, 3696. DOI
10. Chen, Z.; Huang, H.; Wu, J.; Wang, Y. Zero-faulty sample machinery fault detection via relation network with out-of-distribution data augmentation. *Eng. Appl. Artif. Intell.* **2025**, *141*, 109753. DOI
11. Tang, H.; Zhang, K.; Wang, B.; et al. Early bearing fault diagnosis for imbalanced data in offshore wind turbine using improved deep learning based on scaled minimum unscented kalman filter. *Ocean. Eng.* **2024**, *300*, 117392. DOI
12. Chawla, N. V.; Bowyer, K. W.; Hall, L. O.; Kegelmeyer, W. P. SMOTE: synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321-57. DOI
13. Zhao, R.; Yan, R.; Chen, Z.; Mao, K.; Wang, P.; Gao, R. X. Deep learning and its applications to machine health monitoring. *Mech. Syst. Signal. Process.* **2019**, *115*, 213-37. DOI

14. Smith, W. A.; Randall, R. B. Rolling element bearing diagnostics using the Case Western Reserve University data: a benchmark study. *Mech. Syst. Signal. Process.* **2015**, *64–5*, 100–31. DOI
15. Lessmeier, C.; Kimotho, J. K.; Zimmer, D.; Sextro, W. Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: a benchmark data set for data-driven classification. *PHM. Soc. Eur. Conf.* **2016**, *3*. DOI
16. Wang, B.; Lei, Y.; Li, N.; Li, N. A hybrid prognostics approach for estimating remaining useful life of rolling element bearings. *IEEE. Trans. Reliab.* **2020**, *69*, 401–12. DOI
17. Qiu, H.; Lee, J.; Lin, J.; Yu, G. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *J. Sound. Vib.* **2006**, *289*, 1066–90. DOI
18. LaCava, W.; Keller, J.; McNiff, B. Gearbox reliability collaborative: test and model investigation of sun orbit and planet load share in a wind turbine gearbox. In *Proceedings of the 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, Honolulu, USA, Apr 23–26, 2012. DOI
19. Neupane, D.; Seok, J. Bearing fault detection and diagnosis using case Western Reserve University dataset with deep learning approaches: a review. *IEEE. Access.* **2020**, *8*, 93155–78. DOI
20. Huang, M.; Yin, J.; Yan, S.; Xue, P. A fault diagnosis method of bearings based on deep transfer learning. *Simul. Model. Pract. Theory.* **2023**, *122*, 102659. DOI
21. Deng, Y.; Guo, C.; Zhang, Z.; Zou, L.; Liu, X.; Lin, S. An attention-based method for remaining useful life prediction of rotating machinery. *Appl. Sci.* **2023**, *13*, 2622. DOI
22. Nie, L.; Zhang, L.; Xu, S.; Cai, W.; Yang, H. Remaining useful life prediction for rolling bearings based on similarity feature fusion and convolutional neural network. *J. Braz. Soc. Mech. Sci. Eng.* **2022**, *44*, 3638. DOI
23. Zu, H.; Sun, K.; Zhang, W.; Li, C.; Jiao, X. Research on unsupervised fault diagnosis method for wind turbine gearbox. *J. Chin. Soc. Power. Eng.* **2025**, *45*, 106–14. DOI
24. Saito, T.; Rehmsmeier, M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS. One.* **2015**, *10*, e0118432. DOI PubMed PMC
25. Chicco, D.; Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC. Genomics.* **2020**, *21*, 6. DOI PubMed PMC
26. Agost, O.; Aran, F.; Rius, J.; et al. Simet: synthetic image metrics - a synthetic image evaluation framework. *SoftwareX* **2026**, *33*, 102526. DOI
27. Reddi, V. J.; Cheng, C.; Kanter, D.; et al. MLPerf inference benchmark. In *Proceedings of the ACM/IEEE 47th Annual International Symposium on Computer Architecture*. IEEE Press; 2020. pp. 446–59. DOI
28. Fu, W.; Chen, Y.; Li, H.; Chen, X.; Chen, B. Imbalanced fault diagnosis using conditional wasserstein generative adversarial networks with switchable normalization. *IEEE. Sensors. J.* **2023**, *23*, 29119–30. DOI
29. Han, B.; Jia, S.; Liu, G.; Wang, J. Imbalanced fault classification of bearing via wasserstein generative adversarial networks with gradient penalty. *Shock. Vib.* **2020**, *2020*, 1–14. DOI
30. Fathnejat, H.; Nava, V. Data augmentation for damaged scenarios in floating offshore wind turbines: an approach based on diffusion architecture, hierarchical variational approximation and healthy data distribution. *Eng. Comput.* **2025**, *41*, 1979–99. DOI
31. Tang, H.; Gao, S.; Wang, L.; Li, X.; Li, B.; Pang, S. A novel intelligent fault diagnosis method for rolling bearings based on wasserstein generative adversarial network and convolutional neural network under unbalanced dataset. *Sensors* **2021**, *21*, 6754. DOI PubMed PMC
32. Wang, X.; Jiang, H.; Wu, Z.; Yang, Q. Adaptive variational autoencoding generative adversarial networks for rolling bearing fault diagnosis. *Adv. Eng. Inform.* **2023**, *56*, 102027. DOI
33. Wang, C.; Huang, C.; Zhang, L.; et al. Denoising diffusion implicit model combined with TransNet for rolling bearing fault diagnosis under imbalanced data. *Sensors* **2024**, *24*, 8009. DOI PubMed PMC
34. Yang, L.; Zhang, Z.; Song, Y.; et al. Diffusion models: a comprehensive survey of methods and applications. *ACM. Comput. Surv.* **2024**, *56*, 1–39. DOI
35. Wang, X.; Jiang, H.; Zeng, T.; Dong, Y. An adaptive fused domain-cycling variational generative adversarial network for machine fault diagnosis under data scarcity. *Inf. Fusion.* **2026**, *126*, 103616. DOI
36. Hou, Y.; Cai, X.; Chen, Z.; Huang, H.; Zhai, X. A novel fault diagnosis framework based on adaptive VAEGAN and optimal data selection for imbalanced data. *Measurement* **2025**, *256*, 118117. DOI
37. Huang, Y.; Wen, Z.; Wang, F.; Shi, Y. Periodic impact feature enhanced diffusion model for rolling bearing fault diagnosis under small sample conditions. *IEEE. Trans. Instrum. Meas.* **2025**, *74*, 1–14. DOI
38. Feng, Z.; Liang, M.; Chu, F. Recent advances in time–frequency analysis methods for machinery fault diagnosis: a review with application examples. *Mech. Syst. Signal. Process.* **2013**, *38*, 165–205. DOI
39. 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.* **2023**, *219*, 119373. DOI

40. Wang, H.; Yao, L.; Wang, H.; et al. Supervised manifold learning based on multi-feature information discriminative fusion within an adaptive nearest neighbor strategy applied to rolling bearing fault diagnosis. *Sensors* **2023**, *23*, 9820. DOI PubMed PMC
41. Zhou, J.; Zhao, Y. A distance-aware approach for reliable out-of-distribution detection of wind turbine gearbox fault diagnosis. *Front. Energy. Res.* **2024**, *12*, 1496130. DOI
42. Liu, Y.; Sheng, L.; Li, C.; Xi, X.; Orchard, M. E. PhysGAIN: a physical-constraint-based imputation model for missing SCADA data in wind turbines. *IFAC. PapersOnLine.* **2025**, *59*, 7-12. DOI
43. Luo, Y.; Yang, Y.; Kang, S.; Tian, X.; Liu, S.; Sun, F. Wind turbine bearing failure diagnosis using multi-scale feature extraction and residual neural networks with block attention. *Actuators* **2024**, *13*, 401. DOI
44. Turnbull, A.; Carroll, J.; McDonald, A. Combining SCADA and vibration data into a single anomaly detection model to predict wind turbine component failure. *Wind. Energy.* **2021**, *24*, 197-211. DOI
45. Xu, Z.; Yang, P.; Zhao, Z.; Lai, C. S.; Lai, L. L.; Wang, X. Fault diagnosis approach of main drive chain in wind turbine based on data fusion. *Appl. Sci.* **2021**, *11*, 5804. DOI
46. Hassan, I. U.; Panduru, K.; Walsh, J. An in-depth study of vibration sensors for condition monitoring. *Sensors* **2024**, *24*, 740. DOI PubMed PMC
47. Hu, W.; Jiao, Q.; Liu, H.; et al. A transferable diagnosis method with incipient fault detection for a digital twin of wind turbine. *Digit. Eng.* **2024**, *1*, 100001. DOI
48. Li, X.; Wang, Y.; Yao, J.; Li, M.; Gao, Z. Multi-sensor fusion fault diagnosis method of wind turbine bearing based on adaptive convergent viewable neural networks. *Reliab. Eng. Syst. Saf.* **2024**, *245*, 109980. DOI
49. Avendaño-Valencia, L.; Fassois, S. Stationary and non-stationary random vibration modelling and analysis for an operating wind turbine. *Mech. Syst. Signal. Process.* **2014**, *47*, 263-85. DOI
50. Velandia-Cárdenas, C.; Vidal, Y.; Pozo, F.; Casciati, F. Enhancing wind turbine diagnostics with SCADA-vibration fusion, contrastive learning, and linear predictive coefficients. *Struct. Control. Health. Monit.* **2026**, *2026*, 4023580. DOI
51. Parisi, G. I.; Kemker, R.; Part, J. L.; Kanan, C.; Wermter, S. Continual lifelong learning with neural networks: a review. *Neural. Netw.* **2019**, *113*, 54-71. DOI PubMed
52. Khosla, P.; Teterwak, P.; Wang, C.; et al. Supervised contrastive learning. *arXiv* **2020**, arXiv:2004.11362. Available online: <https://doi.org/10.48550/arXiv.2004.11362>. (accessed 2026-09-22).
53. Sohn, K.; Berthelot, D.; Li, C. L.; et al. FixMatch: simplifying semi-supervised learning with consistency and confidence. *arXiv* **2020**, arXiv:2001.07685. Available online: <https://doi.org/10.48550/arXiv.2001.07685>. (accessed 2026-09-22).
54. Sun, S.; Hu, W.; Liu, Y.; Wang, T.; Chu, F. Matching contrastive learning: an effective and intelligent method for wind turbine fault diagnosis with imbalanced SCADA data. *Expert. Syst. Appl.* **2023**, *223*, 119891. DOI
55. Pan, S. J.; Yang, Q. A survey on transfer learning. *IEEE. Trans. Knowl. Data. Eng.* **2010**, *22*, 1345-59. DOI
56. Jiang, P.; Wang, Y.; Wu, S.; Zhang, L.; Yang, C. Fault diagnosis of wind turbine pitch bearings via transfer learning and an improved residual network. *Measurement* **2025**, *254*, 117985. DOI
57. Zhu, Y.; Pei, Y.; Wang, A.; Xie, B.; Qian, Z. A partial domain adaptation scheme based on weighted adversarial nets with improved CBAM for fault diagnosis of wind turbine gearbox. *Eng. Appl. Artif. Intell.* **2023**, *125*, 106674. DOI
58. Li, X.; Gu, J.; Li, M.; et al. Adaptive expert ensembles for fault diagnosis: a graph causal framework addressing distributional shifts. *Mech. Syst. Signal. Process.* **2025**, *234*, 112762. DOI
59. Xu, Y.; Luo, X.; Yuan, M.; Huang, B.; Malof, J. M. Soft-masks guided faster region-based convolutional neural network for domain adaptation in wind turbine detection. *Front. Energy. Res.* **2023**, *10*, 1083005. DOI
60. Cao, L.; Wang, X.; Zhang, H.; Meng, Z.; Li, J.; Liu, M. A novel cross-scenario transferable RUL prediction network with multisource domain meta transfer learning for wind turbine bearings. *IEEE. Trans. Instrum. Meas.* **2025**, *74*, 1-9. DOI
61. Behara, R. K.; Saha, A. K. Comparative performance analysis of deep learning-based diagnostic and predictive models in grid-integrated doubly fed induction generator wind turbines. *Energies* **2025**, *18*, 4725. DOI
62. Lei, Z.; Li, Z. S.; Wen, G.; et al. SFUGDA: source-free unsupervised multiscale graph domain adaptation network with privacy-preserving for cross-domain fault diagnosis of offshore wind turbines. *Mech. Syst. Signal. Process.* **2025**, *235*, 112896. DOI
63. Liu, J.; Cao, H.; Dhupia, J. S.; et al. An adaptive source-free unsupervised domain adaptation method for mechanical fault detection. *Mech. Syst. Signal. Process.* **2025**, *228*, 112475. DOI
64. Yu, A.; Cai, B.; Wu, Q.; García, M. M.; Li, J.; Chen, X. Source-free domain adaptation method for fault diagnosis of rotation machinery under partial information. *Reliab. Eng. Syst. Saf.* **2024**, *248*, 110181. DOI
65. Xiang, S.; Li, P.; Luo, J.; Qin, Y. Micro transfer learning mechanism for cross-domain equipment RUL prediction. *IEEE. Trans. Automat. Sci. Eng.* **2025**, *22*, 1460-70. DOI
66. Qian, Q.; Luo, J.; Qin, Y. Adaptive intermediate class-wise distribution alignment: a universal domain adaptation and generalization method for machine fault diagnosis. *IEEE. Trans. Neural. Netw. Learn. Syst.* **2025**, *36*, 4296-310. DOI

67. Wang, X.; Wang, C.; Liu, H.; et al. An adversarial single-domain generalization network for fault diagnosis of wind turbine gearboxes. *J. Mar. Sci. Eng.* **2023**, *11*, 2384. DOI
68. Guo, P.; Cui, B.; Zhang, W. Cross-device target migration intelligent diagnosis method of wind power gearbox faults. *Meas. Sci. Technol.* **2024**, *35*, 025025. DOI
69. Zhang, Y.; Liu, W.; Gu, H.; Alexisa, A.; Jiang, X. A novel wind turbine fault diagnosis based on deep transfer learning of improved residual network and multi-target data. *Meas. Sci. Technol.* **2022**, *33*, 095007. DOI
70. Zhou, Z.; Chen, W. Entropy-weighted manifold-adjusted transfer learning for cross-condition fault diagnosis with imbalanced and missing labels. *Signal. Process.* **2025**, *229*, 109806. DOI
71. Zhou, H.; Huang, X.; Wen, G.; et al. Convolution enabled transformer via random contrastive regularization for rotating machinery diagnosis under time-varying working conditions. *Mech. Syst. Signal. Process.* **2022**, *173*, 109050. DOI
72. Liu, J.; Wan, L.; Xie, F.; et al. Cross-machine deep subdomain adaptation network for wind turbines fault diagnosis. *Mech. Syst. Signal. Process.* **2024**, *210*, 111151. DOI
73. Liu, X.; Ma, H.; Liu, Y. A novel transfer learning method based on conditional variational generative adversarial networks for fault diagnosis of wind turbine gearboxes under variable working conditions. *Sustainability* **2022**, *14*, 5441. DOI
74. Lu, B.; Dibaj, A.; Gao, Z.; Nejad, A. R.; Zhang, Y. A class-imbalance-aware domain adaptation framework for fault diagnosis of wind turbine drivetrains under different environmental conditions. *Ocean. Eng.* **2024**, *296*, 116902. DOI
75. Cheng, C.; Shan, D.; Teng, Y.; Zhao, B.; Peng, Z.; He, Q. Semisupervised fault diagnosis for gearboxes: a novel method based on a hybrid classification network and weighted pseudo-labeling. *IEEE. Sensors. J.* **2023**, *23*, 16373–83. DOI
76. Wei, C.; Cai, S.; Yang, H.; Zhang, Y. Advanced deep learning-based methodology for multi-class diagnosis of wind turbine blade faults at the wind farm level. *Eng. Appl. Artif. Intell.* **2025**, *160*, 111801. DOI
77. Yue, R.; Jiang, G.; Jin, X.; He, Q.; Xie, P. Spatio-temporal feature alignment transfer learning for cross-turbine blade icing detection of wind turbines. *IEEE. Trans. Instrum. Meas.* **2024**, *73*, 1–17. DOI
78. Zhou, P.; Wang, L.; Wang, Z.; Luo, Z.; Xu, J. Transfer learning-enhanced fault diagnosis for wind turbine main bearings under diverse wind farm conditions. *Insight* **2025**, *67*, 80–9. DOI
79. Karniadakis, G. E.; Kevrekidis, I. G.; Lu, L.; Perdikaris, P.; Wang, S.; Yang, L. Physics-informed machine learning. *Nat. Rev. Phys.* **2021**, *3*, 422–40. DOI
80. Liu, D.; Cui, L.; Cheng, W. Fault diagnosis of wind turbines under nonstationary conditions based on a novel tacho-less generalized demodulation. *Renew. Energy* **2023**, *206*, 645–57. DOI
81. Zhang, Y.; Zhou, Q.; Imani, F.; Tang, J. Seeing the unseen: toward training-free inspection for wind turbine blades using knowledge-augmented vision language models. *IEEE. Access.* **2026**, *14*, 91606–23. DOI
82. Liu, Y.; Cheng, W.; Wen, W.; Fang, Q.; Wu, H. A novel domain generalization framework for fault diagnosis of rotating machinery based on causal representation learning and causal feature identification. *Eng. Appl. Artif. Intell.* **2026**, *164*, 113201. DOI
83. De Lange, M.; Aljundi, R.; Masana, M.; et al. A continual learning survey: defying forgetting in classification tasks. *IEEE. Trans. Pattern. Anal. Mach. Intell.* **2022**, *44*, 3366–85. DOI
84. Li, J.; Yue, K.; Chen, Z.; Xia, J.; Li, W.; Zhang, X. An uncertainty-aware continual learning framework for fault diagnosis of rotating machinery with homogeneous-heterogeneous faults. *IEEE. Trans. Autom. Sci. Eng.* **2026**, *23*, 3284–98. DOI
85. Zong, X.; Yang, R.; Wang, H.; et al. Semi-supervised transfer learning method for bearing fault diagnosis with imbalanced data. *Machines* **2022**, *10*, 515. DOI
86. Guo, C.; Pleiss, G.; Sun, Y.; Weinberger, K. Q. On calibration of modern neural networks. *arXiv* **2017**, arXiv:1706.04599. Available online: <https://doi.org/10.48550/arXiv.1706.04599>. (accessed 2026-09-22).
87. van Engelen, J. E.; Hoos, H. H. A survey on semi-supervised learning. *Mach. Learn.* **2020**, *109*, 373–440. DOI
88. Hüllermeier, E.; Waegeman, W. Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. *Mach. Learn.* **2021**, *110*, 457–506. DOI
89. Han, S.; Sun, S.; Pang, K.; et al. Parallel multi-scale convolutional neural network with incremental learning for bearing compound fault diagnosis under variable working conditions. *J. Braz. Soc. Mech. Sci. Eng.* **2025**, *47*, 5822. DOI
90. Tang, H.; Li, J.; Feng, W.; Chen, P.; Xue, H. Towards machinery incremental fault diagnosis based on inverted transformer lifelong learning with learnable pruning mechanism. *Eng. Appl. Artif. Intell.* **2025**, *152*, 110763. DOI
91. Kirkpatrick, J.; Pascanu, R.; Rabinowitz, N.; et al. Overcoming catastrophic forgetting in neural networks. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 3521–6. DOI PubMed PMC
92. Lu, J.; Lu, S.; Xu, J.; Yan, R. Unknown fault diagnosis of motors based on incremental learning and edge computing. *IEEE. Internet. Things. J.* **2025**, *12*, 22243–56. DOI
93. Ye, Z.; Wu, J.; He, X.; Wang, L.; Jiang, W. Exemplar-free class incremental learning for rotating machinery fault diagnosis via adaptive prototype correction and separation network. *Adv. Eng. Inform.* **2025**, *66*, 103420. DOI

94. Li, X.; Teng, W.; Wang, L.; et al. Trend-constrained pairing based incremental transfer learning for remaining useful life prediction of bearings in wind turbines. *Expert. Syst. Appl.* **2025**, *263*, 125731. [DOI](#)
95. Gama, J.; Žliobaitė, I.; Bifet, A.; Pechenizkiy, M.; Bouchachia, A. A survey on concept drift adaptation. *ACM. Comput. Surv.* **2014**, *46*, 1-37. [DOI](#)
96. Yakupoglu, H.; Gözde, H.; Cengiz Taplamacioglu, M. Online noise-adaptive Kalman filter integrated novel autoencoder for multi-fault detection and early warning of wind turbines. *Measurement* **2025**, *256*, 118538. [DOI](#)
97. Li, H.; Liu, C.; Du, Q.; et al. Sampling-based next-event prediction for wind-turbine maintenance processes. *Energies* **2025**, *18*, 4238. [DOI](#)
98. Shang, X.; Shang, J.; Jiang, Z.; Jiang, C.; Qiu, H.; Gao, L. A multi-task learning approach with meta auxiliary generation network for remaining useful life estimati. *IEEE. Internet. Things. J.* **2025**, *12*, 334-44. [DOI](#)
99. Wang, L.; Bai, Z.; Li, H.; et al. Fault diagnosis of wind turbine pitch bearings based on online soft-label meta-learning and Gaussian prototype network. *Energies* **2025**, *18*, 4437. [DOI](#)
100. Wu, J.; Zhao, Z.; Sun, C.; Yan, R.; Chen, X. Few-shot transfer learning for intelligent fault diagnosis of machine. *Measurement* **2020**, *166*, 108202. [DOI](#)
101. Hendradewa, A. P.; Yin, S. Comparative analysis of offshore wind turbine blade maintenance: RL-based and classical strategies for sustainable approach. *Reliab. Eng. Syst. Saf.* **2025**, *253*, 110477. [DOI](#)
102. Saleh, A.; Chiachío, M.; Salas, J. F.; Kolios, A. Self-adaptive optimized maintenance of offshore wind turbines by intelligent Petri nets. *Reliab. Eng. Syst. Saf.* **2023**, *231*, 109013. [DOI](#)
103. Wang, L.; Cao, H.; Xu, H.; Liu, H. A gated graph convolutional network with multi-sensor signals for remaining useful life prediction. *Knowl. Based. Syst.* **2022**, *252*, 109340. [DOI](#)
104. Zhang, G.; Wang, T.; Chu, F. Mechanism-informed degradation assessment for machinery incorporating monitoring and diagnosis under nonstationary operating conditions. *IEEE/ASME. Trans. Mechatron.* **2026**, *31*, 861-71. [DOI](#)
105. Zhao, R.; Tang, B.; Zhang, X.; Deng, L. Physics-informed cross layer temporal frequency transformer network for remaining useful life prediction of rolling bearings. *Eng. Appl. Artif. Intell.* **2025**, *159*, 111653. [DOI](#)
106. Zhu, Q.; Xiong, Q.; Yang, Z.; Yu, Y. RGCNU: recurrent graph convolutional network with uncertainty estimation for remaining useful life prediction. *IEEE/CAA. J. Autom. Sinica.* **2023**, *10*, 1640-2. [DOI](#)
107. Yucesan, Y. A.; Viana, F. A. Physics-informed digital twin for wind turbine main bearing fatigue: quantifying uncertainty in grease degradation. *Appl. Soft. Comput.* **2023**, *149*, 110921. [DOI](#)
108. Hu, C.; Goebel, K.; Howey, D.; et al. Editorial: Special issue on physics-informed machine learning enabling fault feature extraction and robust failure prognosis. *Mech. Syst. Signal. Process.* **2023**, *192*, 110219. [DOI](#)
109. Zhu, J.; Pei, T.; Su, L.; Lan, B.; Chen, W. Blended ensemble learning for robust normal behavior modeling of wind turbines. *Energy. Sci. Eng.* **2025**, *13*, 2565-84. [DOI](#)
110. Stetco, A.; Dinmohammadi, F.; Zhao, X.; et al. Machine learning methods for wind turbine condition monitoring: a review. *Renew. Energy.* **2019**, *133*, 620-35. [DOI](#)
111. Zhang, W.; Vatn, J.; Rasheed, A. Gearbox pump failure prognostics in offshore wind turbine by an integrated data-driven model. *Appl. Energy.* **2025**, *380*, 124829. [DOI](#)
112. He, S.; Hou, W.; Chen, Z.; Liu, H.; Wang, J.; Cheng, P. Early warning model based on support vector machine ensemble algorithm. *J. Oper. Res. Soc.* **2025**, *76*, 411-25. [DOI](#)
113. Liu, Z.; He, X.; Huang, B.; Zhou, D. Incremental learning-enabled fault diagnosis of dynamic systems: a comprehensive review. *IEEE. Trans. Cybern.* **2025**, *55*, 5633-49. [DOI PubMed](#)
114. Guo, Y.; Zhi, Y.; Qi, R.; Cai, M. A lightweight 1D-CNN-transformer for bearing fault diagnosis under limited data and AWGN interference. *Sensors* **2026**, *26*, 2574. [DOI PubMed PMC](#)
115. Qiao, L.; Zhang, Y.; Wang, Q.; Li, D.; Peng, S. Fault diagnosis for wind turbine generators based on model-agnostic meta-learning: a few-shot learning method. *Expert. Syst. Appl.* **2025**, *267*, 126171. [DOI](#)
116. Coronel, D.; Guevara, C.; Santos, M. Hybrid graph neural network with FFT and genetic optimization for fault detection in wind turbines. *Neurocomputing* **2026**, *684*, 133352. [DOI](#)
117. Tian, J.; Yu, Y.; Karimi, H. R.; Gao, F.; Lin, J. A continual test-time domain adaptation method for online machinery fault diagnosis under dynamic operating conditions. *Neural. Netw.* **2026**, *194*, 108192. [DOI](#)
118. Barber, S.; Lima, L. A. M.; Sakagami, Y.; et al. Enabling co-innovation for a successful digital transformation in wind energy using a new digital ecosystem and a fault detection case study. *Energies* **2022**, *15*, 5638. [DOI](#)
119. de Novaes Pires Leite, G.; Farias, F. C.; de Sá, T. G.; et al. A robust fleet-based anomaly detection framework applied to wind turbine vibration data. *Eng. Appl. Artif. Intell.* **2023**, *126*, 106859. [DOI](#)

120. Raymond, T.; Charbonnier, S.; Bérenguer, C.; Lebranchu, A. Fleet-wide data-driven model generation process for residual-based fault detection in wind energy. *Wind. Energy*. **2025**, *28*, e70050. DOI
121. Jenkel, L.; Jonas, S.; Meyer, A. Privacy-preserving fleet-wide learning of wind turbine conditions with federated learning. *Energies* **2023**, *16*, 6377. DOI
122. Zhou, R.; Li, Y.; Lin, X. A clustered federated learning framework for collaborative fault diagnosis of wind turbines. *Appl. Energy*. **2025**, *377*, 124532. DOI
123. Cheng, X.; Shi, F.; Liu, Y.; Liu, X.; Huang, L. Wind turbine blade icing detection: a federated learning approach. *Energy* **2022**, *254*, 124441. DOI
124. Kan, X.; Chen, X.; Li, L.; Peng, X.; Zhong, M. A fault diagnosis method based on dual-layer adaptive personalized federated learning in wind turbines. *IFAC. PapersOnLine*. **2025**, *59*, 1848–53. DOI
125. Zhou, Z.; Chen, X.; Li, E.; Zeng, L.; Luo, K.; Zhang, J. Edge intelligence: paving the last mile of artificial intelligence with edge computing. *Proc. IEEE*. **2019**, *107*, 1738–62. DOI
126. Dubey, A. K.; Kumar, A.; Segovia Ramírez, I.; García Márquez, F. P. Machine learning and hybrid intelligence for wind energy optimization: a comprehensive state-of-the-art review. *Expert. Syst. Appl.* **2026**, *296*, 128926. DOI
127. Shi, W.; Cao, J.; Zhang, Q.; Li, Y.; Xu, L. Edge computing: vision and challenges. *IEEE. Internet. Things. J.* **2016**, *3*, 637–46. DOI
128. Zhang, W.; Ma, X. Simultaneous fault detection and sensor selection for condition monitoring of wind turbines. *Energies* **2016**, *9*, 280. DOI
129. Hancer, E.; Xue, B.; Zhang, M.; Karaboga, D.; Akay, B. Pareto front feature selection based on artificial bee colony optimization. *Inf. Sci.* **2018**, *422*, 462–79. DOI
130. Xue, B.; Zhang, M.; Browne, W. N.; Yao, X. A survey on evolutionary computation approaches to feature selection. *IEEE. Trans. Evol. Comput.* **2016**, *20*, 606–26. DOI
131. Han, F.; Chen, W.; Ling, Q.; Han, H. Multi-objective particle swarm optimization with adaptive strategies for feature selection. *Swarm. Evol. Comput.* **2021**, *62*, 100847. DOI
132. Dutta, P.; Kanti Podder, K.; Islam Sumon, M. S.; et al. GearFaultNet: novel network for automatic and early detection of gearbox faults. *IEEE. Access.* **2024**, *12*, 188755–65. DOI
133. Zhu, Y.; Liu, X. A lightweight CNN for wind turbine blade defect detection based on spectrograms. *Machines* **2023**, *11*, 99. DOI
134. Zhao, H.; Gao, Y.; Deng, W. Defect detection using shuffle Net-CA-SSD lightweight network for turbine blades in IoT. *IEEE. Internet. Things. J.* **2024**, *11*, 32804–12. DOI
135. Wu, Y.; Tang, B.; Deng, L.; Shen, Y. Hardware-resource-constrained neural architecture search for edge-side fault diagnosis of wind-turbine gearboxes. *IEEE. Trans. Ind. Electron.* **2024**, *71*, 9812–22. DOI
136. Cao, X.; Gao, Z.; Lu, H.; Tang, J.; Yao, Y.; Liu, L. WCFormer: a wavelet-enhanced CNN-transformer hybrid network for bearing fault diagnosis using multi-sensor signal fusion. *J. Comput. Des. Eng.* **2025**, *12*, 82–100. DOI
137. Yoo, Y.; Jo, H.; Ban, S. W. Lite and efficient deep learning model for bearing fault diagnosis using the CWRU dataset. *Sensors* **2023**, *23*, 3157. DOI PubMed PMC
138. Mo, C.; Huang, K.; Ji, H. A lightweight and precision dual track 1D and 2D feature fusion convolutional network for machinery equipment fault diagnosis. *Sci. Rep.* **2024**, *14*, 31666. DOI PubMed PMC
139. Cheng, X.; Dou, S.; Du, Y.; Wang, Z. Gearbox fault diagnosis method based on lightweight channel attention mechanism and transfer learning. *Sci. Rep.* **2024**, *14*, 743. DOI PubMed PMC
140. Sze, V.; Chen, Y.; Yang, T.; Emer, J. S. Efficient processing of deep neural networks: a tutorial and survey. *Proc. IEEE*. **2017**, *105*, 2295–329. DOI
141. Astolfi, D.; De Caro, F.; Vaccaro, A. Condition monitoring of wind turbine systems by explainable artificial intelligence techniques. *Sensors* **2023**, *23*, 5376. DOI PubMed PMC
142. Floreale, G.; Baraldi, P.; Lu, X.; Rossetti, P.; Zio, E. Sensitivity analysis by differential importance measure for unsupervised fault diagnostics. *Reliab. Eng. Syst. Saf.* **2024**, *243*, 109846. DOI
143. Chatterjee, J.; Dethlefs, N. Deep learning with knowledge transfer for explainable anomaly prediction in wind turbines. *Wind. Energy*. **2020**, *23*, 1693–710. DOI
144. Liu, X.; Cui, S.; Qiao, W.; Liu, J.; Wu, G. Remaining useful life prediction with uncertainty quantification for rotating machinery: a method based on explainable variational deep Gaussian process. *Neurocomputing* **2025**, *638*, 130232. DOI
145. Borgonovo, E.; Apostolakis, G. A new importance measure for risk-informed decision making. *Reliab. Eng. Syst. Saf.* **2001**, *72*, 193–212. DOI
146. Rajaoarisoa, L.; Randrianandraina, R.; Nalepa, G. J.; Gama, J. Decision-making systems improvement based on explainable artificial intelligence approaches for predictive maintenance. *Eng. Appl. Artif. Intell.* **2025**, *139*, 109601. DOI

147. Yao, Q.; Chen, B.; Hu, A.; Zhen, D.; Xiang, L. A novel diffusion model with Shapley value analysis for anomaly detection and identification of wind turbine. *Expert. Syst. Appl.* **2025**, *284*, 127925. DOI
148. Han, Y.; Lv, S.; Huang, Q.; Zhang, Y. AMCW-DFNSA: an interpretable deep feature fusion network for noise-robust machinery fault diagnosis. *Knowl. Based. Syst.* **2024**, *301*, 112361. DOI
149. Huang, T.; Zhang, Q.; Lu, X.; Zhao, S.; Yang, S. Integrating outlier-type prior knowledge into convolutional neural networks based on an attention mechanism for fault diagnosis. *IEEE. Trans. Syst. Man. Cybern. Syst.* **2024**, *54*, 7834-47. DOI
150. Lv, S.; Huang, Q.; Cui, T.; Zhang, Y.; Han, Y. AGMFormer: an explainable intelligent diagnosis framework based on adaptive Gaussian mixture constrained feature aggregation network. *IEEE. Trans. Instrum. Meas.* **2025**, *74*, 1-12. DOI
151. Oliveira-Filho, A.; Zemouri, R.; Cambron, P.; Tahan, A. Early detection and diagnosis of wind turbine abnormal conditions using an interpretable supervised variational autoencoder model. *Energies* **2023**, *16*, 4544. DOI
152. Yang, T.; Jiang, L.; Guo, Y.; Han, Q.; Li, X. LTFM-net framework: advanced intelligent diagnostics and interpretability of insulated bearing faults in offshore wind turbines under complex operational conditions. *Ocean. Eng.* **2024**, *309*, 118533. DOI
153. Yang, T.; Xu, M.; Chen, C.; Wen, J.; Li, J.; Han, Q. DSTF-Net: a novel framework for intelligent diagnosis of insulated bearings in wind turbines with multi-source data and its interpretability. *Renew. Energy.* **2025**, *238*, 121965. DOI
154. Berghout, T.; Benbouzid, M. UBO-EREX: uncertainty Bayesian-optimized extreme recurrent EXpansion for degradation assessment of wind turbine bearings. *Electronics* **2024**, *13*, 2419. DOI
155. Cao, L.; Zhang, H.; Meng, Z.; Wang, X. A parallel GRU with dual-stage attention mechanism model integrating uncertainty quantification for probabilistic RUL prediction of wind turbine bearings. *Reliab. Eng. Syst. Saf.* **2023**, *235*, 109197. DOI
156. García-Vaca, M.; Sierra-García, J.; Santos, M. Probabilistic evaluation for early wind turbine yaw misalignment detection. *Reliab. Eng. Syst. Saf.* **2026**, *266*, 111716. DOI
157. Jing, H.; Zhao, C. Adjustable piecewise regression strategy based wind turbine power forecasting for probabilistic condition monitoring. *Sustain. Energy. Technol. Assess.* **2022**, *52*, 102013. DOI
158. Li, X.; Teng, W.; Zhang, Y.; Peng, D.; Liu, Y. Dynamic normalized health indicator construction and Bayesian recurrent state estimation for remaining useful life prediction of high-speed bearings in wind turbine drivetrain. *Measurement* **2025**, *246*, 116725. DOI
159. Adedipe, T.; Shafiee, M.; Zio, E. Bayesian network modelling for the wind energy industry: an overview. *Reliab. Eng. Syst. Saf.* **2020**, *202*, 107053. DOI
160. Sun, Y.; Yang, X.; Hu, S.; Sun, L.; Kang, J. Failure analysis of floating offshore wind turbines considering working condition correlations. *Qual. Reliab. Eng.* **2023**, *39*, 1889-908. DOI
161. Yang, T.; Li, G.; Huo, J.; Li, X.; Zhou, X.; Han, Q. The DRTS-Net framework: enhanced probabilistic health perception for wind turbine bearings under complex operating conditions with improved interpretability. *IEEE. Trans. Instrum. Meas.* **2025**, *74*, 1-12. DOI
162. Zheng, Y.; Bi, R.; Peng, Y. Non-Gaussian deep latent-variable state-space model fusing domain knowledge. *IEEE. Sensors. J.* **2025**, *25*, 2352-61. DOI
163. Zhu, Y.; Zhao, S.; Zhang, Y.; Zhang, C.; Wu, J. A review of statistical-based fault detection and diagnosis with probabilistic models. *Symmetry* **2024**, *16*, 455. DOI
164. Xiao, Y.; Feng, K.; Miao, D.; Zhang, P.; Yang, J. A state-supervised model and novel anomaly index for gas turbines blade fault detection under multi-operating conditions. *IEEE. Access.* **2025**, *13*, 14225-38. DOI
165. Zhang, Z.; Ouyang, T.; Shen, X. Risk-regularization optimization for learning wind power curve. *IEEE. Trans. Emerg. Top. Comput. Intell.* **2025**, 1-8. DOI
166. Amin, A.; Bibo, A.; Panyam, M.; Tallapragada, P. An uncertainty-aware health monitoring model for wind turbine drivetrains based on Bayesian neural network. *IFAC. PapersOnLine.* **2023**, *56*, 235-40. DOI
167. Pandit, R.; Infield, D. Gaussian process operational curves for wind turbine condition monitoring. *Energies* **2018**, *11*, 1631. DOI
168. Zhao, R. An improved set-valued observer and probability density function-based self-organizing neural networks for early fault diagnosis in wind energy conversion systems. *Symmetry* **2025**, *17*, 448. DOI
169. Movsessian, A.; Cava, D. G.; Tcherniak, D. Interpretable machine learning in damage detection using Shapley Additive Explanations. *ASCE. ASME. J. Risk. Uncertain. Eng. Syst. B. Mech. Eng.* **2022**, *8*, 021101. DOI
170. Xu, Q.; Jiang, H.; Zhang, X.; Li, J.; Chen, L. Multiscale convolutional neural network based on channel space attention for gearbox compound fault diagnosis. *Sensors* **2023**, *23*, 3827. DOI PubMed PMC
171. Cui, S.; Liu, X.; Wu, G.; Qiao, W.; Liu, J.; Chen, X. Accurate remaining useful life prediction and uncertainty quantification for wind turbines using temporal convolutional variational deep Gaussian processes. *Sci. Rep.* **2026**, *16*, 16791. DOI PubMed PMC
172. Ozdemir, A. A.; Seiler, P.; Balas, G. J. Wind turbine fault detection using counter-based residual thresholding. *IFAC. Proc. Vol.* **2011**, *44*, 8289-94. DOI
173. Raissi, M.; Perdikaris, P.; Karniadakis, G. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **2019**, *378*, 686-707. DOI

174. Nie, Y.; Lu, H.; Chen, X.; et al. Interpreting vibration characteristics of planetary gearboxes for condition monitoring based on a novel transfer path effect model. *Nonlinear. Dyn.* **2025**, *113*, 9351–71. DOI
175. Guidotti, R.; Monreale, A.; Ruggieri, S.; Turini, F.; Giannotti, F.; Pedreschi, D. A survey of methods for explaining black box models. *ACM. Comput. Surv.* **2019**, *51*, 1–42. DOI
176. Scholkopf, B.; Locatello, F.; Bauer, S.; et al. Toward causal representation learning. *Proc. IEEE.* **2021**, *109*, 612–34. DOI
177. Oh, S. Y.; Joung, C.; Lee, S.; et al. Condition-based maintenance of wind turbine structures: a state-of-the-art review. *Renew. Sustain. Energy. Rev.* **2024**, *204*, 114799. DOI
178. Peng, H.; Zhang, H.; Fan, Y.; Shangguan, L.; Yang, Y. A review of research on wind turbine bearings' failure analysis and fault diagnosis. *Lubricants* **2023**, *11*, 14. DOI
179. Peng, H.; Zhang, H.; Shangguan, L.; Fan, Y. Review of tribological failure analysis and lubrication technology research of wind power bearings. *Polymers* **2022**, *14*, 3041. DOI PubMed PMC
180. Peng, H.; Zhao, D.; Shangguan, L.; Li, S.; Cheng, R. Review of wind power bearing wear analysis and intelligent lubrication method research. *Coatings* **2024**, *14*, 30. DOI
181. Qiu, W.; Zhao, X.; Tyrrell, A.; Perinpanayagam, S.; Niu, S.; Wen, G. Application of artificial intelligence-based technique in electric motors: a review. *IEEE. Trans. Power. Electron.* **2024**, *39*, 13543–68. DOI
182. Yang, Y.; Liang, F.; Zhu, Q.; Zhang, H. An overview on structural health monitoring and fault diagnosis of offshore wind turbine support structures. *J. Mar. Sci. Eng.* **2024**, *12*, 377. DOI
183. Li, M.; Xu, Z.; Li, S.; et al. Health prognostics and maintenance decision-making for wind energy: a comprehensive overview. *Renew. Sustain. Energy. Rev.* **2026**, *226*, 116269. DOI
184. Li, W.; Xiang, D.; Liang, F. Multi-objective opportunity maintenance optimization for wind farms considering multiple factors. *Renew. Energy.* **2026**, *256*, 123868. DOI
185. Rozas, H.; Xie, W.; Gebraeel, N. Condition-based maintenance for wind farms using a distributionally robust chance constrained program. *IEEE. Trans. Power. Syst.* **2025**, *40*, 231–43. DOI
186. Lee, N.; Woo, J.; Kim, S. A deep reinforcement learning ensemble for maintenance scheduling in offshore wind farms. *Appl. Energy.* **2025**, *377*, 124431. DOI
187. Zhang, W.; Vatn, J.; Rasheed, A. A differential evolution based approach for short-term predictive maintenance optimization of an offshore wind turbine. *Ocean. Eng.* **2025**, *335*, 121507. DOI
188. Ziegler, L.; Gonzalez, E.; Rubert, T.; Smolka, U.; Melero, J. J. Lifetime extension of onshore wind turbines: a review covering Germany, Spain, Denmark, and the UK. *Renew. Sustain. Energy. Rev.* **2018**, *82*, 1261–71. DOI
189. Szubartowski, M.; Migawa, K.; Borowski, S.; Neubauer, A.; Hujo, L.; Kopiláková, B. Application of the semi-Markov processes to model the Enercon E82-2 preventive wind turbine maintenance system. *Energies* **2024**, *17*, 199. DOI
190. Tao, Z.; Zhu, R.; Hu, J.; Wang, M.; Chen, Q.; Wang, C. A novel hierarchical failure analysis approach targeting the operation and maintenance of floating offshore wind turbines. *Renew. Energy.* **2025**, *241*, 122267. DOI
191. Frederiksen, R. D.; Bocewicz, G.; Nielsen, P.; Radzki, G.; Banaszak, Z. A reference modelling approach for cost optimal maintenance for offshore wind farms. *Sustainability* **2024**, *16*, 8352. DOI
192. El-Barbary, Z. M. S.; Safarova, L.; Atamurotov, F.; Alsayah, A. M.; Yadav, B. K.; Saha, A. K. Artificial intelligence driven Internet of Things framework for wind energy monitoring and performance enhancement in smart cities. *Int. J. Energy. Res.* **2025**, *2025*, 7215655. DOI
193. Song, Z.; Wang, X.; Wu, Q.; et al. A task offloading strategy based on multi-agent deep reinforcement learning for offshore wind farm scenarios. *Comput. Mater. Contin.* **2024**, *81*, 985–1008. DOI
194. Radiuk, P.; Rusyn, B.; Melnychenko, O.; et al. Criticality assessment of wind turbine defects via multispectral UAV fusion and fuzzy logic. *Energies* **2025**, *18*, 4523. DOI
195. Megahed, T. F.; Morgan, E. F.; Timo, P. N.; Saeed, M. Neural network predictive control for fault detection and identification in DFIG with SMES for low voltage ride-through requirements. *Ain. Shams. Eng. J.* **2024**, *15*, 102775. DOI
196. Nasir, M.; Maiti, A.; Xu, S.; Garg, S. Controller modelling for situation awareness in industrial internet of things: current trends and opportunities with artificial intelligence. *Int. J. Syst. Sci.* **2026**, *57*, 2495–530. DOI
197. Ghods, M.; Tabarniarami, Z.; Faiz, J. Real-time detection of short circuit faults of wind turbine permanent magnet synchronous generator by short-time analysis of the signal amplitude ripple intensity based on analytical model. *IEEE. Trans. Power. Delivery.* **2024**, *39*, 1588–99. DOI
198. Huang, H.; Li, Z.; Huang, S.; Zhang, Z. A fast open-circuit fault diagnosis method of parallel wind-turbine converters via zero-sequence circulating current informed residual analysis. *Energies* **2025**, *18*, 4801. DOI
199. He, Y.; Liu, J.; Wu, S.; Wang, X. Condition monitoring and fault detection of wind turbine driveline with the implementation of deep residual long short-term memory network. *IEEE. Sensors. J.* **2023**, *23*, 13360–76. DOI

200. Song, C.; Zhang, N.; Shao, J.; Wang, Y.; Liu, X.; Jiang, C. Diagnosis of wind turbine yaw system based on self-attention–long short-term memory (LSTM). *Electronics* **2025**, *14*, 617. DOI
201. Liu, J.; Wang, X.; Xie, F.; Wu, S.; Li, D. Condition monitoring of wind turbines with the implementation of spatio-temporal graph neural network. *Eng. Appl. Artif. Intell.* **2023**, *121*, 106000. DOI
202. Lyu, Q.; Liu, J.; He, Y.; Wang, X.; Wu, S. Condition monitoring of wind turbines with implementation of interactive spatio-temporal deep learning networks. *IEEE. Trans. Instrum. Meas.* **2023**, *72*, 1-11. DOI
203. Wang, X.; Zheng, Z.; Jiang, G.; He, Q.; Xie, P. Detecting wind turbine blade icing with a multiscale long short-term memory network. *Energies* **2022**, *15*, 2864. DOI
204. Miao, Q.; Wang, D.; Xia, Z.; Xu, C.; Zhan, J.; Wu, C. Exploring spatio-temporal dynamics for enhanced wind turbine condition monitoring. *Mech. Syst. Signal. Process.* **2025**, *223*, 111841. DOI
205. Qian, X.; Sun, T.; Zhang, Y.; Wang, B.; Awad Gendeel, M. A. Wind turbine fault detection based on spatial-temporal feature and neighbor operation state. *Renew. Energy.* **2023**, *219*, 119419. DOI
206. Xu, Z.; Ou, H.; Bashir, M.; Beer, M.; Wang, J.; Yang, Z. Dynamic reliability framework for resilience assessment for floating offshore wind turbines. *J. Reliab. Sci. Eng.* **2026**, *2*, 035501. DOI
207. Bisoyi, S.; Kumar Rathi, A.; Mahato, S. Reliability assessment of remaining useful life predictions of roller bearings using deep learning models. *J. Reliab. Sci. Eng.* **2026**, *2*, 025302. DOI
208. Gao, M.; Gao, J.; Niu, Y.; Wang, H.; He, Y.; Li, Y. Reliability analysis of wind turbine with feed-forward structure based on GO-FLOW method. *J. Reliab. Sci. Eng.* **2026**, *2*, 035001. DOI
209. Schlögl, M.; Schneider, M.; Schmid, U. Scalable electromagnetic energy harvester for wind turbine rotor blade applications. *Smart. Mater. Struct.* **2024**, *33*, 055037. DOI
210. Masabi, S. N.; Fu, H.; Flint, J. A.; Theodossiades, S. A pendulum-based rotational energy harvester for self-powered monitoring of rotating systems in the era of industrial digitization. *Appl. Energy.* **2024**, *365*, 123200. DOI
211. Zheng, M.; Man, J.; Wang, D.; Chen, Y.; Li, Q.; Liu, Y. Semi-supervised multivariate time series anomaly detection for wind turbines using generator SCADA data. *Reliab. Eng. Syst. Saf.* **2023**, *235*, 109235. DOI
212. Zhu, L.; Song, Y. Sequential data-driven automatic calibration of wind turbine fault information in smart grids. *IEEE. Internet. Things. J.* **2022**, *9*, 8943-52. DOI
213. Sharma, S.; Nava, V. Condition monitoring of mooring systems for floating offshore wind turbines using convolutional neural network framework coupled with autoregressive coefficients. *Ocean. Eng.* **2024**, *302*, 117650. DOI
214. Liu, X.; Du, J.; Ye, Z. A condition monitoring and fault isolation system for wind turbine based on SCADA data. *IEEE. Trans. Ind. Inform.* **2022**, *18*, 986-95. DOI

Disclaimer/Publisher's Note: All statements, opinions, and data contained in this publication are solely those of the individual author(s) and contributor(s) and do not necessarily reflect those of OAE and/or the editor(s). OAE and/or the editor(s) disclaim any responsibility for harm to persons or property resulting from the use of any ideas, methods, instructions, or products mentioned in the content.



© The Author(s) 2026. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.