Article citation info:
Guan M, Wan L, Liu W. Cross-domain diagnosis model for early fault of wind turbine drive system based on digital twin and federated learning. Diagnostyka. 2026;27(3):2026303. https://doi.org/10.29354/diag/225499.

DIAGNOSTYKA, 2026, Vol. 27, No. 3

e-ISSN 2449-5220
DOI: 10.29354/diag/225499

CROSS-DOMAIN DIAGNOSIS MODEL FOR EARLY FAULT OF WIND TURBINE DRIVE SYSTEM BASED ON DIGITAL TWIN AND FEDERATED LEARNING

Meng GUAN [1,*], Li WAN [2], Weiwei LIU [3]

[1] Mechanical Engineering School of Shenyang Institute of Engineering, Shenyang 110136, Liaoning, China
[2] Neutech Group Limited, Shenyang 110000, Liaoning, China
[3] Xiamen Intelligent-Benefit Bank Science and Technology Limited, Xiamen 361000, Fujian, China
[*] Corresponding author, e-mail: guanmeng_mail@163.com

Abstract

A cross-domain diagnosis approach combining Digital Twin (DT) and Federated Learning (FL) is proposed to further optimize the issues of sample scarcity and privacy protection in early defect diagnosis of wind turbine transmission systems. Firstly, a DT which combines both mechanism- and data-driven approaches is proposed to generate multi-condition virtual fault data. Next, the feature alignment strategy is used to reduce the domain difference between simulated and real data. Finally, a cross-wind FL framework is proposed, which uses dual-path feature learning and dynamic aggregation strategy combined with data quality evaluation for synchronous training. The experimental results on Coverage, Accuracy, Reliability and Earliness (CARE) dataset show that: 1) the average F1 score of the proposed DT-FL model on three wind farm test sets is 0.942, which is better than that of the FL (F1 score is 0.907) and the standard federated average algorithm (F1 score is 0.925) which only use measured data. 2) In the cross-wind migration task, the average performance attenuation of DT-FL model is 7.6%, and it has better generalization ability. The above results show that the proposed model can effectively use virtual data to enhance the diagnostic capabilities and improve the early fault identification performance under the premise of protecting data privacy.

Keywords: wind turbine; transmission system; early fault diagnosis; digital twin; federated learning

1. INTRODUCTION

To guarantee the dependability and cost-effectiveness of wind power, wind turbines must operate steadily over the long term, and their transmission systems are prone to malfunctions [1,2]. Due to strong noise interference and normal operational vibrations, the weak early fault signal characteristics are easily concealed, which brings difficulties to the accurate diagnosis of fans [3]. Additionally, during the actual operation of wind farms, the number of available fault samples for training is limited, and the operating conditions, unit models, and data distributions of each wind farm differ significantly, complicating model construction [4]. At the same time, due to factors such as commercial competition and data security, various wind farms are usually reluctant to share local monitoring data, which leads to the occlusion of data circulation, thus making it impossible to conduct centralized data training [5, 6]. To protect user privacy and safety, it is crucial to investigate a problem diagnosis technique for wind turbines using small samples and diverse structural data distributions.

At present, Digital Twin (DT) technology provides a foundation for the establishment of high-precision virtual simulation models and can effectively compensate for the lack of actual fault samples by generating virtual fault data [7]. However, the method of training solely with simulation data may cause the feature distribution of virtual data to deviate from the actual observations, thus affecting the diagnostic performance [8]. Furthermore, joint learning is a distributed machine learning technique that, in theory, allows each participant to train the model locally, better meeting customers' privacy protection demands [9]. However, there are variations in the quantity and quality of local data from various users in the wind power scenario. A small number of fault features are difficult for the global Federated Learning (FL) aggregation model to train efficiently, which impacts the model's overall performance and fairness [10].

This study suggests a cross-domain diagnosis model that combines DT and FL to address the above issues. The goal of the model is to provide each wind farm client with a DT that combines multi-body dynamics mechanisms and data-driven calibration to generate a virtual early fault sample containing

Received 2026-01-07; Accepted 2026-07-06; Available online 2026-07-21

multiple operating conditions. In addition, a FL framework across wind fields and a dynamic weighted optimization algorithm, which combines the local data quality and model verification performance of clients, are proposed to integrate knowledge from different clients. In summary, the goal of this study is to enhance the accuracy of early defect diagnosis in wind turbine systems.

2. RELATED WORKS

The current study on intelligent fault detection of wind turbine transmission systems mostly uses high-fidelity simulation modeling, data-driven methods, and privacy protection techniques. For instance, Jiang et al. developed a FL framework for cloud-side collaboration that reduced communication costs and efficiently utilized distributed data to increase problem identification efficiency [11]. Based on the data characteristics of offshore wind farms, Lu et al. proposed an event-triggered FL method that guaranteed diagnostic accuracy while lowering communication frequency and cost [12]. Tan et al. proposed a fault detection framework for hydropower stations, which combined DT and deep learning, enhancing the model's sensitivity to early anomalies in a complex operating environment [13]. Cheng et al. proposed a cluster FL model using blockchain technology. The model recorded the training process through the blockchain, ensuring the traceability of data exchange and the safety of model updates [14]. Fera and Spandonidis proposed a DT framework based on Artificial Intelligence (AI). It was verified that this method had certain potential for improving diagnostic accuracy and system visualization analysis [15]. In their discussion of condition monitoring enhancement for wind turbine gearboxes based on DT and AI, Habbouche et al. noted that DT could offer richer physical information contexts, boosting the monitoring model's capacity for generalization [16]. In their discussion of generative AI's role in developing a DT for fault detection in predictive maintenance, Mikołajewska et al. demonstrated how it could efficiently provide a variety of fault data to support model training [17]. Rana summarized the application of AI in power system fault detection and predictive maintenance, including data-driven methods, DT, and self-healing power grids. He also pointed out that the integration of digital models and distributed learning was an important trend for intelligent operation and maintenance in the future [18].

To sum up, although the existing research has made some achievements in using FL to protect data privacy and using DT to generate data, how to build a diagnosis model with rich sample size and safety needs to be studied.

3. CROSS-DOMAIN DIAGNOSIS MODEL COMBINING DT AND FL

3.1. Method of Virtual Data Generation and Feature Enhancement Based on DT

There are now issues with early wind field defect identification, such as a lack of samples and a variety of data formats that are difficult to combine [19]. This study suggests a framework that combines DT and FL to address these issues. Fig. 1 depicts the framework's workflow.

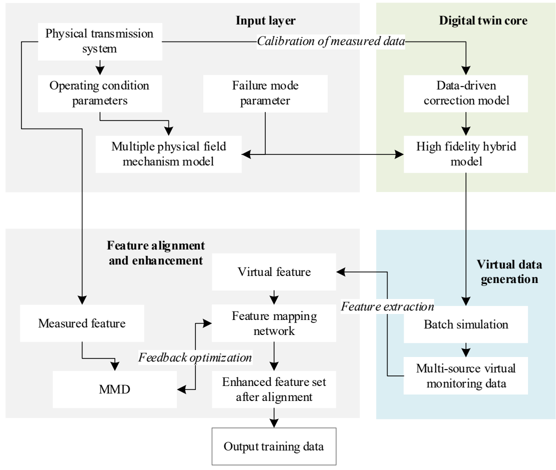

Fig. 1. Virtual data generation and feature enhancement process based on DT

The spindle, gearbox, and generator bearing are the main mechanical elements of the wind turbine's transmission system that are highlighted in Fig. 1 [20]. Among them, the DT is composed of a multi-physical field mechanism model and a data-driven model. The construction of the mechanism model is based on multi-body dynamics and load analysis, which is responsible for simulating the response of the transmission chain under normal working conditions and different fault states [21]. Its key dynamic behavior is expressed by the forced vibration equation:

$$M\ddot{x}(t) + C\dot{x}(t) + Wx(t) = F(t, \theta, \xi) \quad (1).$$

M, C and W represent the mass, damping and stiffness matrices of the system respectively. $x(t)$ is the displacement response vector of the system. $F(t, \theta, \xi)$ represents the excitation force vector, which is a function composed of time t, operating condition parameter θ and fault state parameter ξ. By adjusting ξ, the specified failure mode can be simulated.

The data-driven model uses limited field normal operation data, which is used to calibrate the parameters and correct the errors of the mechanism model to ensure that the twin output is consistent with the behavior of the physical entity in a healthy state [22]. The previously mentioned way of merging mechanism and data seeks to secure the physical interpretability of the model and improve the proximity of the model to the real system's reaction.

In the above-mentioned high-fidelity DT, a multi-condition virtual fault sample generation strategy is also adopted, with the aim of including all kinds of operating conditions and early fault forms that may be encountered in the actual wind field. However, if the virtual simulation data is directly used to train the diagnosis model, the "simulation-to-reality" domain difference problem may arise. Therefore, to avoid this situation, the feature alignment and enhancement mechanism of virtual and measured data is added to the model [23]. The basic principle of this mechanism is to extract and strengthen the fault-sensitive and cross-domain stable feature representations. Firstly, the virtual signal and the available field-measured signal are extracted in the same time-frequency domain. Then, using the concept of domain adaptation, a feature mapping network G_f is used to transform the virtual features, so that their distribution is aligned with the measured feature distribution. Finally, the maximum mean variance Maximum Mean Discrepancy (MMD) is used to minimize the process. The following is how MMD is calculated:

$$L_{MMD} = \left\| \frac{1}{N_s} \sum_{i=1}^{N_s} \phi(G_f(x_i^s)) - \frac{1}{N_t} \sum_{j=1}^{N_t} \phi(x_j^t) \right\|_H^2 \quad (2)$$

x_i^s and x_j^t represent the feature samples from the virtual source domain and the measured target domain, respectively. N_s and N_t are the corresponding number of samples, and $\phi(\cdot)$ is a function of mapping features to reproducing kernel Hilbert space H. By optimizing the loss, the feature distribution of the generated virtual data will be close to the measured data distribution. Finally, the aligned virtual features and the limited measured features together form an enhanced feature training set, which provides an information basis for the subsequent model training.

3.2. Construction and Optimization of FL Model for Cross-domain Diagnosis

3.2.1. Framework Design of Distributed Diagnosis across Wind Fields

When early fault diagnosis is carried out for the transmission system of wind turbines, the privacy risk of each wind field's data cannot be ignored [24]. Based on model training with local data (including the virtual data generated in the previous section), a distributed framework of cross-wind field FL for early fault diagnosis is proposed. The principle of this framework is to keep the original data in each wind farm at all times, and only exchange the encrypted model parameters to jointly train a globally shared intelligent diagnosis model. Fig. 2 presents the overall architecture and workflow of the distributed federated diagnosis framework across wind farms.

This framework uses the client-server architecture and connects it with the real-world scenario of wind power maintenance and operation in Fig. 2. Two types of crucial data are stored, and each participating wind farm is considered a FL client with full local model training capabilities. After feature extraction, the first is the feature set of the wind turbine's historical and current monitoring data. The other is the virtual fault feature data generated by the DT method proposed in section 3.1.

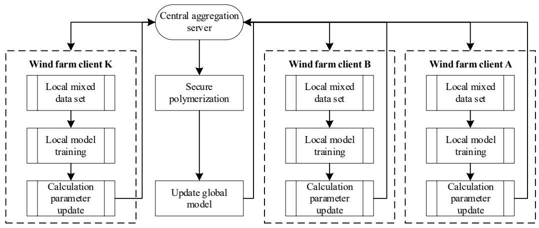

Fig. 2. Overall framework and workflow of distributed federal diagnosis framework across wind farms

2.2. Federated Feature Learning Based on Virtual Data Fusion

Based on the above cross-wind federated framework, a feature learning mechanism integrating virtual data is proposed. The main principle of this mechanism is to use a dual-path feature learning network and a joint loss function, allowing the model to simultaneously extract stable and transportable fault feature representations from both measured data and virtual data [25].

The local diagnosis model of each client is composed of a shared feature encoder $E(\cdot;\theta_e)$ and two classifiers$C_r(\cdot;\theta_r)$ and $C_v(\cdot;\theta_v)$. Among them, E is responsible for extracting high-level abstract representation from input features, and θ_e is a parameter. C_r is used to process the feature representation from the measured data to classify the real faults. C_v is responsible for processing feature representation from virtual data. In each round of local training, for a batch of multi-type data containing N_r measured samples and N_v virtual samples, their feature representations are first obtained by an encoder:

$$z_i^r = E(x_i^r;\theta_e), \quad (3)$$

$$z_j^v = E(x_j^v;\theta_e) \quad (4)$$

x_i^r and x_j^v represent the feature inputs of the *i*-th measured sample and the *j*-th virtual sample, respectively. z_i^r and z_j^v are the corresponding encoded feature vectors.

To make better use of virtual data and prevent it from interfering with the distribution of real data, the classification loss of the two kinds of data will be calculated separately. The classification loss of measured data adopts standard cross entropy loss:

$$L_{cls}^r = -\frac{1}{N_r} \sum_{i=1}^{N_r} \sum_{c=1}^{C} y_{i,c}^r \log(\hat{y}_{i,c}^r) \quad (5)$$

C is the total number of fault categories, $y_{i,c}^r$ is the indication value of the real label of the *i*-th measured sample on category c, and $\hat{y}_{i,c}^r = C_r(z_i^r)$ is the prediction probability of the model that the sample belongs to category c. The classification loss L_{cls}^v of virtual data is calculated in the same way, but the classifier C_v and the virtual sample label y^v are used.

Similarly, the alignment loss of feature distribution is adopted to ensure that the encoder can learn the features beneficial to real fault diagnosis from virtual data. This loss is achieved by calculating a simple second-order moment matching loss:

$$\mathcal{L}_{align} = \left\| \frac{1}{N_r}\sum_{i=1}^{N_r} z_i^r - \frac{1}{N_v}\sum_{j=1}^{N_v} z_j^v \right\|_2^2 + \|\mathrm{Cov}(Z^r) - \mathrm{Cov}(Z^v)\|_F^2 \quad (6)$$

The first term is responsible for making the mean values of the two types of features as close as possible, and the second term is to make their covariance matrices $\mathrm{Cov}(Z^r)$ and $\mathrm{Cov}(Z^v)$ as similar as possible. $\|\cdot\|_F$ stands for Frobenius norm. The minimization of this function shows that the encoder will be guided to learn a common feature space, so that the fault mode knowledge contained in virtual data can better help the measured data to complete the classification.

After the above operations, the local training joint loss function is obtained by the weighted sum method:

$$L_{local} = L_{cls}^r + \alpha L_{cls}^v + \beta L_{align} \quad (7)$$

α and β are hyperparameters, which are used to control the weights of virtual data classification loss and feature alignment loss. In each round of local training, the client minimizes the function by random gradient descent algorithm, and updates its local model parameters $\Theta = \{\theta_e, \theta_r, \theta_v\}$.

3.2.3. Optimization of Federated Aggregation Combining Different Types of Data

The central server must combine models from several clients to update during the FL aggregation stage. Based on this, this section proposes a federated aggregation optimization algorithm combined with data quality evaluation.

According to the model update situation uploaded by different clients, this variable is used to measure the ratio of real and effective fault samples in the local data set of clients. If a client has more high-quality measured fault data, its trained local model update is generally more reliable. Therefore, a weighting factor based on data composition is given first. Assuming that the number of measured fault samples is N_k^{real} and the number of virtual fault samples is N_k^{virt} in the local data set of the *k*th client, the total data of the client is the sum of the above two. At this point, a ratio is used to describe the true weight of its data:

$$w_k^{base} = \frac{N_k^{real}}{N_k^{real} + \gamma \cdot N_k^{virt}} \quad (8)$$

γ is a discount factor ranging from 0 to 1, which is used to adjust the contribution of virtual data. However, the quality of the update cannot be fully evaluated only by the data ratio. Therefore, the performance of a client local model on the local verification set is constructed as another evaluation index. Assume that the accuracy of the verification set of the *k*-th client local model in the current round is A_k. Then the obtained aggregation weight λ_k is calculated by the data base weight and the performance weight:

$$\lambda_k = \frac{w_k^{base} \cdot \exp(A_k/\tau)}{\sum_{j=1}^{K}(w_j^{base} \cdot \exp(A_j/\tau))} \quad (9)$$

τ is a temperature parameter used to smooth the weight distribution, and K is the total number of clients participating in the aggregation. Clients with high data quality and strong local model performance will be given more weight in this equation.

This study optimizes the traditional federal average aggregation after providing the dynamic weight of each customer. The server performs weighted aggregation after receiving each client's model parameter update $\Delta\theta_k$, which is the difference between the local model parameters and the most recent round of global parameters.

$$\Delta\Theta_g = \sum_{k=1}^{K} \lambda_k \cdot \Delta\theta_k \quad (10)$$

$$\Theta_g^{t+1} = \Theta_g^t + \eta \cdot \Delta\Theta_g \quad (11)$$

$\Delta\Theta_g$ is the calculated global model parameters whose update directions Θ_g^t and Θ_g^{t+1} are the t and $t+1$ rounds respectively, and η is the global learning rate.

3.2.4. Model Optimization of Early Fault Sensitivity

The defect characteristic signal in wind turbine transmission system early fault diagnostics is typically very faint and readily obscured by significant background noise and normal operational vibrations. To increase the model's diagnostic sensitivity and make it focus more on the important features that are susceptible to early errors, this part will employ an attention mechanism.

The channel attention module is added to the local feature encoder E of the client. The goal of this module is to make the model automatically learn the importance weights of different characteristic channels in early fault diagnosis. It is assumed that the middle feature map output by the encoder is $U \in \mathrm{R}^{m \times l \times p}$, where p is the number of channels and m and l are the spatial dimensions of the feature map. Firstly, U is globally averaged and pooled to obtain the global description of each channel:

$$s_p = \frac{1}{m \times l}\sum_{i=1}^{m}\sum_{j=1}^{l} U_p(i,j) \quad (12)$$

s_p is the global statistic of the p-th channel. $U_p(i,j)$ is the eigenvalue of the p-th channel at position (i,j). Then, through a simple two-layer neural network and nonlinear activation function, the final channel attention weight vector $a = [a_1, a_2, \dots, a_p]$ is generated. The weight vector is normalized to (0,1) by Sigmoid function. Finally, the feature map U is recalibrated by its corresponding attention weight:

$$\widetilde{U_p} = a_p \cdot U_p \quad (13)$$

This process enables the model to dynamically enhance the characteristic channels containing early fault information, while suppressing irrelevant and noisy channels, thus amplifying the performance of weak fault signals in the characteristic space.

4. EXPERIMENTAL RESULTS AND ANALYSIS

4.1. Experimental Setup and Data Description

(1) Selection of dataset

In this experiment, the dataset Coverage, Accuracy, Reliability and Earliness (CARE) is selected for the follow-up study. This dataset contains high-quality operation data from three different wind farms (36 turbines in total), including many monitored variables, including generator speed, output power, multi-directional vibration signals of gearbox bearings, etc. (dataset link: https://doi.org/10.5281/zenodo.14006163). To simulate the FL scene across wind farms, 36 wind turbines from three different wind farms are naturally divided into three federated clients, and the data of each client represents the local data distribution of an independent wind farm. The data division technique is displayed in Table 1, where each client's local data is further separated into training set (60%), verification set (15%), and test set (25%) in chronological sequence.

Table 1. Division of federal dataset based on wind field

Federated client (wind farm)	Number of fans	Number of normal behavior periods	Number of marked early abnormal periods
Client A	12	18	16
Client B	12	17	14
Client C	12	16	14

(2) Experimental platform

The specific configuration of the experimental platform used in this study is shown in Table 2, which includes a complete environment from hardware infrastructure to software framework and key library versions.

(3) Model parameter setting

The hyperparameter settings related to model training and FL are determined by previous ablation experiments, and the specific values are shown in Table 3.

4.2. Model Validity Verification

4.2.1. The Performance Analysis of Early Fault Diagnosis

The basic ability of DT-FL model in early fault identification is verified by comparative experiments. Baseline models include independent Local-Only training model, classical federal average model Federal Average (FedAvg) and Centralized training model (this is a theoretical model, assuming that all wind field data can be centralized in one center and all measured data can be used for training). Accuracy, precision, recall, and F1-score are examples of evaluation indicators. Figs. 3 and 4 show the performance indicators of each model in three client test sets and the category of "early failure".

Table 2. Details of experimental environment configuration

Type	Configuration project	Detailed description
Hardware	Computing server	Rack server
	CPU	2 Intel Xeon Gold 5218(32 cores /64 threads)
	RAM	128 GB DDR4
	GPU	2 NVIDIA T4(16GB GDDR6 memory)
	Save	1TB NVMe SSD+4TB SATA HDD
Software and framework	Operating system	Ubuntu 20.04 LTS
	Programming language	Python 3.8
	Deep learning framework	PyTorch 1.12.0+CUDA 11.6
	FL simulation	Custom FL framework based on PyTorch
	Scientific calculation and simulation	MATLAB R2022a, Simulink
	Other key libraries	NumPy, SciPy, scikit-learn, Pandas

Table 3. Model parameter setting

Parameter category	Parameter name	Setting value
Federal training	Number of clients	3
	Total number of communication rounds	100
	Number of local training rounds	5
	synchronous frequency	1
	Client sampling rate	1.0
	Aggregation weight discount factor	0.3
Local model optimization	Batch size	64
	Initial learning rate	0.001
	Optimizer	Adam
	Virtual data loss weight	0.7
	Feature alignment loss weight	0.5
	Focus loss parameter	2.0
	Temperature parameter	0.5

Fig. 3 and Fig. 4 show that DT-FL model has better performance in early fault identification of wind farm. Compared with Local-Only model (F1 score is 0.853) and FedAvg model (F1 score is 0.900), the F1 score of DT-FL model in three wind fields is 0.942. The main reason for the above phenomenon is that the DT-FL model generates virtual fault samples under multiple operating conditions using digital twin technology, which effectively supplements the missing early fault samples in each wind farm, enables the model to perceive more fault modes, and thereby improves the recognition accuracy of real faults.

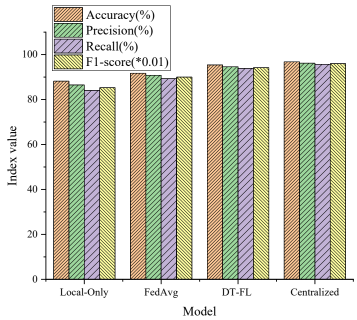

Fig. 3. Comparison of average performance of different diagnostic models on various wind field test sets

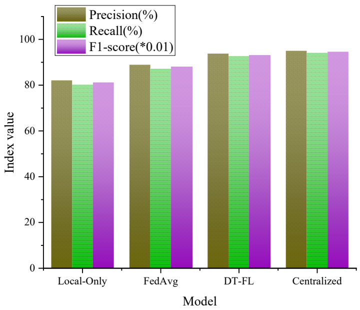

Fig. 4. Diagnostic performance of each model for "early fault" category

On the identification of early fault categories, the recall rate (92.5%) and F1 score (0.931) of DT-FL are higher than those of Local-Only by 12.2% and 0.119, respectively. This is mainly attributed to the introduction of the channel attention mechanism in the model, which enables the neural network to focus more on weak frequency-band components related to early faults.

This demonstrates that by using DT technology to introduce virtual data, DT-FL can successfully address the issue of data heterogeneity and enhance the detection capacity of a few types of faults.

4.2.2. Generalization Ability Test Across Different Scenarios

A cross-wind migration experiment is set up to test the generalization ability. One wind field is selected as the target domain, and its data is only used for testing. The model is trained using the data from the other two wind fields as the source domain. Three training methods were set up in the experiment, including centralized training in the source domain (X1), federated training in the source domain (X2), and fine-tuning in the target domain (X3). In addition, the average performance attenuation of each model when it is directly migrated to different target domains is also counted. Performance degradation is defined as the difference between the model's F1 score in the target domain and its average F1 score in the source domain. Finally, the results shown in Figs. 5 and 6 are obtained.

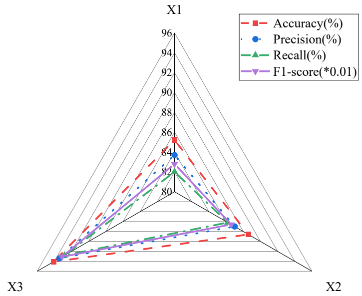

Fig. 5. Cross-wind field generalization performance test (target domain: wind field C)

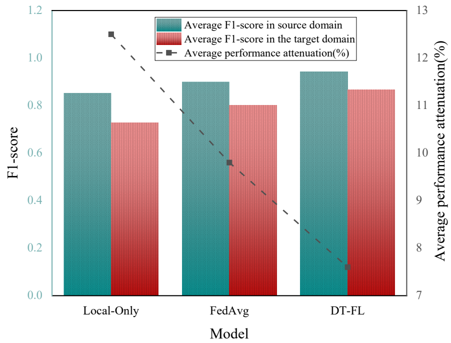

Fig. 6. Average performance attenuation of direct migration across wind fields of different models

Fig. 5 shows that the DT-FL framework is 0.038 higher than the centralized training in the source domain (0.828) only through the federal training in the source domain (F1 score is 0.866) when the wind field C is taken as the target domain. After fine-tuning with 10% target domain data, the F1 score is raised to 0.931. This is because the vibration data of wind farm C in the target domain exhibits working condition drift, and a small amount of local data can be used to correct the error between the virtual features and actual features of the digital twin model, resulting in improved performance.

Fig. 6 shows that the average attenuation rate of DT-FL (7.6%) is lower than that of Local-Only (12.5%) and FedAvg (9.8%). It shows that its DT virtual data strategy can effectively alleviate the negative migration problem caused by data heterogeneity. This result occurs because the digital twin takes into account different operating conditions (wind speed, load, etc.) during virtual data generation, and the features it obtains are less affected by varying wind farm conditions. Consequently, the performance does not degrade significantly when transferred to new wind farms.

4.2.3. Key Component Ablation Experiment

In this study, ablation experiments are carried out, and three model variants were constructed: 1) Remove DT virtual data and using only measured data for FL (W/O DT); 2) Keep the virtual data, but use the standard federal average algorithm to aggregate and remove the dynamic weighted optimization (W/OOPT); 3) Complete DT-Full Model. Table 4 displays the final experimental results.

Table 4. Experimental analysis of ablation of key components

Model variant	Accuracy (%)	Precision (%)	Recall rate (%)	F1 score
Full Model	95.4	94.6	93.9	0.942
W/o DT (measured federal only)	92.1	91.0	90.5	0.907
W/o Opt (standard polymerization)	93.8	92.9	92.0	0.925

From Table 4, the F1 score decreases by 0.035 due to the removal of dummy data, which shows that dummy data plays an important role in expanding samples and enhancing early fault characteristics. Removing aggregation optimization reduces F1 score by 0.017. This demonstrates how the dynamic weighting technique based on data quality and model performance can more successfully integrate the expertise of various clients, hence raising the global model's quality.

4.2.4. Model Parameters and Sensitivity Analysis

In this study, the sensitivity of virtual data loss weight α to model performance is analyzed. Other parameters in the experiment are fixed, and when α changes from 0.1 to 1.0, the model's performance on the verification set changes. Fig. 7 shows this change.

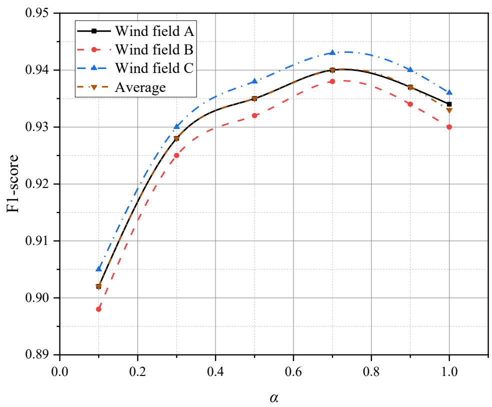

Fig. 7. Sensitivity analysis of virtual data loss weight α (F1 score of verification set)

Fig. 7 illustrates that when α is too little (0.1), the model performs mediocrely and is unable to fully utilize the information in virtual data. With the increase of α, the performance gradually increases and reaches the maximum when α=0.7. When α continues to increase to 1.0, the performance decreases, which shows that over-reliance on virtual data may slightly interfere with learning the real data distribution. Therefore, α=0.7 is a better value, which can make the model make full use of the extended knowledge in virtual data. This phenomenon can be interpreted as follows: although virtual data contain abundant fault features, they still exhibit certain deviations compared with real data. Excessive weighting will cause the model to generate spurious correlations, thereby reducing its recognition performance for actual faults.

4.3. Discussion

The proposed DT-FL model achieves favorable early fault identification performance on the CARE dataset. To further understand the merits and demerits of the presented method, this study compares it with two recent relevant works. One is the cloud-edge collaborative federated learning framework proposed in [11], which mainly reduces communication costs and improves computational efficiency but does not incorporate a virtual data generation strategy. The other is the event-triggered federated learning strategy proposed in [12], which mainly addresses the data transmission frequency issue in offshore wind farms and has limited capability in processing different types of data. Different from these two studies, this study introduces dual-path feature learning into the federated learning framework, enabling virtual data and real data to participate in training jointly, while employing a feature alignment method to narrow the distribution discrepancy between them. In addition, during the fusion process, it proposes a dynamic weight assignment method based on data quality and local validation performance. The advantage of using this method is that clients with high data quality can contribute more to the update of the global model, thereby reducing model bias caused by irrelevant and heterogeneous data. Overall, the research idea of this study differs from existing studies, yet its application in complex operating environments and multi-fault scenarios remains to be further verified.

5. CONCLUSION

To address the issues of data scarcity, significant structural variations among wind farms, and privacy protection in the early fault diagnosis of wind turbine transmission systems, this study proposes a cross-domain diagnosis model that combines DT and FL. On the one hand, the study builds a DT that creates multi-condition virtual fault samples by combining data-driven and multi-physical field mechanisms. On the other hand, the feature alignment approach is employed to reduce the domain gap between measured and simulation data. Furthermore, dual-path feature learning and a dynamic weighted aggregation technique are incorporated into a cross-wind-field FL framework to enhance fault identification accuracy and generalization.

Experiments show that the model is superior to traditional local training and FedAvg methods on CARE datasets, and it also performs better in early fault detection. However, the limitation of the study is that the model is only verified for specific types of transmission parts and limited wind field data, and extreme working conditions and compound fault scenarios are not considered. Future research will expand the model to include more fault types and complex operating environments, further improving its engineering practicability and real-time diagnosis ability.

Source of funding: *This work was supported by Liaoning Provincial Department of Education Basic Research Project (No. JYTMS20230302).*

Author contributions: *Research concept and design, M.G., L.W.; Collection and/or assembly of data, M.G.; Data analysis and interpretation, M.G., L.W.; Writing the article, M.G., L.W., W.L.; Critical revision of the article M.G.; final approval of the article, M.G., W.L.*

Declaration of competing interest: *The authors declares no conflict of interest.*

REFERENCES

1. Peng H, Zhang H, Fan Y, et al. A review of research on wind turbine bearings' failure analysis and fault diagnosis. Lubricants. 2022;11(1):14. https://doi.org/10.3390/lubricants11010014.
2. Gu H, Liu W Y, Gao Q W, et al. A review on wind turbines gearbox fault diagnosis methods. Journal of Vibroengineering. 2021;23(1):26-43. https://doi.org/10.21595/jve.2020.20178.
3. Zhao Y, Song Z, Li D, et al. Wind turbine gearbox fault diagnosis based on multi-sensor signals fusion. Protection and Control of Modern Power Systems. Protection and Control of Modern Power Systems, 2024;9(4):96-109. https://doi.org/10.23919/PCMP.2023.000241.
4. Gao Z, Liu X. An overview on fault diagnosis, prognosis and resilient control for wind turbine systems. Processes. 2021;(2):300. https://doi.org/10.3390/pr9020300.
5. Yang Q, Tang B, Shen Y, et al. Self-attention parallel fusion network for wind turbine gearboxes fault diagnosis. IEEE Sensors Journal. 2023;23(19): 23210-23220. https://doi.org/10.1109/JSEN.2023.3308971.
6. Amin A, Bibo A, Panyam M, et al. Wind turbine gearbox fault diagnosis using cyclostationary analysis and interpretable CNN. Journal of Vibration Engineering & Technologies. 2024;12(2):1695-1705. https://doi.org/10.1007/s42417-023-00937-1.
7. Stergiou CL, Psannis KE. Digital twin intelligent system for industrial internet of things-based big data management and analysis in cloud environments. Virtual Reality & Intelligent Hardware. 2022;4(4): 279-291. https://doi.org/10.1016/j.vrih.2022.05.003.
8. Singh M, Fuenmayor E, Hinchy EP, et al. Digital twin: Origin to future. applied system innovation. 2021;4(2):36. https://doi.org/10.3390/asi4020036.
9. Thilakarathne NN, Muneeswari G, Parthasarathy V, et al. Federated learning for privacy-preserved medical internet of things. Intell. Autom. Soft Comput. 2022; 33(1):157-172. https://doi.org/10.32604/iasc.2022.023763.
10. Hotz VJ, Bollinger CR, Komarova T, et al. Balancing data privacy and usability in the federal statistical system. Proceedings of the National Academy of Sciences. 2022;119(31):e2104906119. https://doi.org/10.1073/pnas.2104906119.
11. Jiang G, Zhao K, Liu X, et al. A federated learning framework for cloud–edge collaborative fault diagnosis of wind turbines. IEEE Internet of Things Journal. 2024;11(13):23170-23185. https://doi.org/10.1109/JIOT.2024.3387417.
12. Lu S, Gao Z, Zhang P, et al. Event-triggered federated learning for fault diagnosis of offshore wind turbines with decentralized data. IEEE Transactions on Automation Science and Engineering. 2023;21(2): 1271-1283. https://doi.org/10.1109/TASE.2023.3270354.
13. Tan J, Radhi R M, Shirini K, et al. Innovative framework for fault detection and system resilience in hydropower operations using digital twins and deep learning. Scientific Reports. 2025;15(1):15669. https://doi.org/10.1038/s41598-025-98235-1.
14. Cheng X, Tian W, Shi F, et al. A blockchain-empowered cluster-based federated learning model for blade icing estimation on IoT-enabled wind turbine. IEEE Transactions on Industrial Informatics. 2022; 18(12):9184-9195. https://doi.org/10.1109/TII.2022.3159684.
15. Fera F, Spandonidis C. A fault diagnosis approach utilizing artificial intelligence for maritime power systems within an integrated digital twin framework. Applied Sciences. 2024;14(18):8107. https://doi.org/10.3390/app14188107.
16. Habbouche H, Amirat Y, Benbouzid M. Leveraging digital twins and AI for enhanced gearbox condition monitoring in wind turbines: A Review. Applied Sciences. 2025;15(10):5725. https://doi.org/10.3390/app15105725.
17. Mikołajewska E, Mikołajewski D, Mikołajczyk T, et al. Generative AI in AI-based digital twins for fault diagnosis for predictive maintenance in Industry 4.0/5.0]. Applied Sciences. 202515(6):3166. https://doi.org/10.3390/app15063166.
18. Rana S. AI-driven fault detection and predictive maintenance in electrical power systems: A systematic review of data-driven approaches, digital twins, and self-healing grids. American Journal of Advanced Technology and Engineering Solutions. 2025;1(01): 258-289. https://doi.org/10.63125/4p25x993.
19. Mukherjee M, Batabyal S, Roy S D, et al. Deep learning-based fault detection and classification in power distribution networks. Lex Localis. 2025; 23(S6):1915-1928. https://doi.org/10.52152/vdwhe059.
20. Mammadov A, Danilov Y. Predictive maintenance and digital twins for greener power generation: Case studies from China, Germany, Norway, and The Netherlands. European Journal of Energy Research. 2025;5(6):12-17. https://doi.org/10.11648/j.ijepe.20251405.11.
21. Xu Z, Yang P, Zhao Z, et al. Fault diagnosis approach of main drive chain in wind turbine based on data fusion. Applied Sciences. 2021;11(13):5804. https://doi.org/10.3390/app11135804.
22. Teng W, Ding X, Tang S, et al. Vibration analysis for fault detection of wind turbine drivetrains – a

comprehensive investigation. Sensors. 2021;21(5): 1686. https://doi.org/10.3390/s21051686.
23. Liang G, Hu J, Yang K, et al. Data augmentation for predictive digital twin channel: Learning multi-domain correlations by convolutional TimeGAN. IEEE Journal of Selected Topics in Signal Processing. 2024;18(1):18-33. https://doi.org/10.1109/JSTSP.2024.3358980.
24. Chen H, Hsu H Y, Hsieh J Y, et al. A differential privacy-preserving federated learning scheme with predictive maintenance of wind turbines based on deep learning for feature compression and anomaly detection with state assessment. Journal of Mechanical Science and Technology. 2024;38(7):3413-3429. https://doi.org/10.1007/s12206-024-0616-9.
25. Xiao Y, Shao H, Lin J, et al. BCE-FL: A secure and privacy-preserving federated learning system for device fault diagnosis under non-IID condition in IIoT. IEEE Internet of Things Journal. 2023;11(8):14241-14252. https://doi.org/10.1109/JIOT.2023.3340745.

Meng GUAN was born in Shenyang, Liaoning, China, in 1988. She obtained a doctor's degree from Shenyang Agricultural University in China. She is currently working at the School of mechanical engineering, Shenyang Institute of Engineering. Her main research direction is Fault diagnosis of wind power equipment.
e-mail: guanmeng_mail@163.com

Li WAN was born in Tieling, Liaoning, China, in 1989. She obtained a master's degree from Shenyang Agricultural University in China. She is currently working at Neutech Group Limited. Her main research direction is the application of digital twin technology in mechanical fault diagnosis.
e-mail: wl_mail2025@163.com

Weiwei LIU was born in Anshan, Liaoning, China, in 1990. She obtained a master's degree from Shenyang Agricultural University in China. She is currently working at Xiamen intelligent-benefit bank Science and Technology Limited. Her main research direction is design and maintenance of wind power equipment.
e-mail: monica89271698@163.com