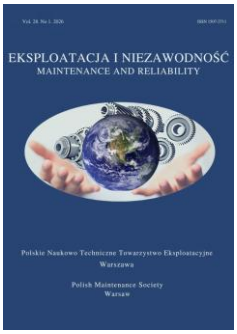

Eksploatacja i Niezawodnosc – Maintenance and Reliability

Volume 28 (2026), Issue 1

journal homepage: http://www.ein.org.pl

Article citation info:
Liu Q, Yan D, Lu B, Sun Z, Li L, Algorithms for Monitoring Large and Deep Wear Failures of Yaw Bearing Races in Wind Turbines, Eksploatacja i Niezawodnosc – Maintenance and Reliability 2026: 28(1)http://doi.org/10.17531/ein/207795

Algorithms for Monitoring Large and Deep Wear Failures of Yaw Bearing Races in Wind Turbines

Indexed by: Web of Science Group

Qidong Liu[a,*], Dexin Yan[a], Biao Lu[a], Zhiyuan Sun[a], Lin Li[a]

[a]Qinghai Huanghe Wind Power Co., Ltd, China

Highlights

- Missing data & wavelet denoising ensure reliable preprocessing.
- Subdomain adversarial network enables cross-condition fault classification.
- Wear severity scoring with multi-level warnings for proactive maintenance.
- High-accuracy monitoring for complex internal raceway wear.
- Vibration-based method outperforms traditional visual inspections.

Abstract

The yaw bearing of wind turbine is large in size and complex in structure, and the raceway is located inside the bearing, so the traditional monitoring methods, such as the machine vision inspection method, or the use of some auxiliary tools, such as endoscopy, etc., can only carry out local inspection, and they can't set a reasonable reference threshold for faults and categorize the fault characteristics, which results in the inaccuracy of detecting the internal wear faults of the raceway. For this reason, an algorithm is proposed to monitor the large-area deep wear faults of the yaw bearing raceway of wind turbine. The accelerometer (vibration sensor) is selected to collect the bearing state data, and the missing value interpolation is implemented to complete the bearing state data; then the data collected is denoised by the improved wavelet modal maxima denoising algorithm; Based on this, the joint sub-domain adaptive adversarial migration network establishes a classification model of the bearing fault feature extraction, extracts the wear and tear fault features of the bearing, and constructs a diagnostic model to classify and recognize the extracted features; At the same time, according to the degree of wear of the bearing to set the wear severity level, establish early warning mechanism, and after the classification of the failure to carry out the wear level scoring, to find out the corresponding wear level, to complete the monitoring and warning. The experimental results show that when the algorithm is utilized to complete the bearing wear monitoring, the denoising effect of the bearing state data is obvious, and the monitoring accuracy is very high.

Keywords

wind turbine, yaw bearing, raceway wear, fault monitoring, improved wavelet mode maxima denoising algorithm, adversarial migration network

1. Introduction

Under the background of global energy structure transformation and sustainable development strategy, wind power, as a clean and renewable form of energy, is being popularized and applied worldwide at an unprecedented speed [1-2]. As one of the core components of wind power generation system, the yaw bearing bears the key tasks of supporting the wind wheel, transmitting torque and realizing the adjustment of the wind wheel with the wind direction, and its

(*) Corresponding author.
E-mail addresses: Q. Liu (ORCID: 0009-0005-5656-7561) liuqidong23213@163.com, D. Yan (ORCID: 0009-0005-9794-6074) yandexin2312312@163.com, B. Lu (ORCID: 0009-0003-8624-2325) lubiao43541531@163.com, Z. Sun (ORCID: 0009-0005-3569-6460) sunzhiyuan313123@163.com, L. Li (ORCID: 0009-0007-4793-0001) lilin231312312@163.com

performance and reliability are directly related to the overall efficiency, operation life and maintenance cost of the wind turbine. However, with the continuous expansion of wind farm scale and the prolongation of operation time, the wear problem of yaw bearing, especially the large area deep wear in the raceway area, has become an important bottleneck restricting the further development of wind power generation technology and the improvement of economic benefits. The wear problem of yaw bearings has significant particularities compared with main shaft bearings or gearbox bearings, mainly reflected in their dynamic load characteristics and operating conditions. Yaw bearings need to withstand intermittent impact loads caused by frequent start-stop cycles, as well as dynamic yaw moments caused by asymmetric wind loads. This alternating stress will accelerate the contact fatigue on the raceway surface. Meanwhile, the low-speed and heavy-load characteristics of the yaw system make it difficult for the lubricating film to form stably, resulting in a combined damage mode of adhesive wear and abrasive wear on the raceway. In addition, the intermittent fretting wear of the yaw bearing (caused by the swing of the wind wheel) will further aggravate the spalling at the raceway edge, which is essentially different from the continuous rotational wear of the main shaft bearing or the high-frequency meshing wear of the gearbox bearing.

The working environment of yaw bearings is extremely harsh, not only to withstand the extreme weather conditions of the huge load and complex stress, but also need to face the sand, dust, salt spray and other corrosive substances, these factors together accelerate the fatigue and wear process of bearing materials [3-4]. Raceway as the most vulnerable parts of the bearing, the degree of wear is directly related to the bearing's rotational accuracy, vibration characteristics and service life. Once the raceway is deeply worn over a large area, it will not only lead to bearing failure, but also trigger chain reactions, such as gearbox damage, main shaft bending, etc., which will seriously affect the stability and safety of the wind turbine and increase the unplanned downtime and maintenance costs.

In summary, the wear of wind turbine yaw bearing raceway not only affects the bearing capacity and service life of the yaw bearing, but also may lead to the increase of vibration and noise of the whole wind turbine, and even cause serious mechanical failure, resulting in huge economic losses. Therefore, real-time monitoring and diagnosis of the large-area deep wear of wind turbine yaw bearing raceway is of great significance to improve the stability and reliability of wind turbines and reduce the operation and maintenance cost.

The diagnostic method based on motor current signal analysis (MCSA) in the literature [5] is unable to diagnose the minor faults of the rolling bearing outer raceway, and there is the problem of diagnostic failure under no-load or light-load operation conditions of the motor. In this regard, we propose a diagnostic method for the fault diagnosis of asynchronous motor rolling bearing outer raceway based on the rotational frequency multimodulation components of the stator current signal. Firstly, multiple frequency band modulation components are determined according to the rotational frequency; then the stator current signal is analyzed by continuous refinement Fourier transform near each rotational frequency modulation component; the fault index is determined according to the amplitude of multiple rotational frequency modulation components under normal and fault conditions; finally, the fault index is used to determine whether the fault occurs or not to realize the fault monitoring. Due to the serious redundancy of the algorithm parameters in the Fourier transform of the current signal, the method has a low performance in fault monitoring.

Literature [6] method firstly analyzes the acoustic vibration of the equipment unit, and improves the singular value decomposition denoising method by fusing the similar soft thresholds, so as to effectively eliminate the inherent noise interference of the non-contact sensors; Secondly, it proposes the inverse Barker spectral transform method, and extracts the acoustic and vibration feature maps of the unit bearings by combining the Barker spectral transform with the Gram's angle and the field transform and other characteristic engineering techniques; by fusing the relative position encoding self By integrating the self-attention mechanism of relative position encoding and depth-separable convolution, a feature map transfer network is established; At the same time, a time-series data transfer network is constructed by using the multi-head self-attention mechanism and bi-directional long- and short-term memory network, and a parallel grid

architecture is used to construct a fault diagnosis model of the unit bearings, and the bearing state data is collected and inputted into the model, and then based on the output of the model training, the existence of faults is monitored, and the types of faults are determined, so as to complete the monitoring of the wear faults. The method neglects the interference caused by the external influencing factors of bearing operation when extracting the bearing acoustic pattern characteristics, so the method has lower monitoring accuracy when monitoring bearing wear faults.

Literature [7] method firstly adopts the overlapping sampling method to enhance the one-dimensional time series data; then uses the Markov transformation field method to convert the one-dimensional time series data into two-dimensional image representation; then constructs a snake network for wear monitoring and introduces the migration learning technique to optimize the network structure; Finally, inputs the two-dimensional images into a neural network model to provide the two-dimensional image samples and retains the time domain information, and then builds and trains the ResNeXt and ResNeSt improved residual networks based on the migration learning fine-tuning process to classify the fault images and realize the fault diagnosis, and complete the wear monitoring. Due to the mismatch between the migration learning parameters and the network structure when the migration learning principle is introduced, the method is ineffective in bearing wear monitoring.

Literature [8] method first established a forced lubrication rolling bearing rotor test bench, artificial introduction of high hardness non-ferrous metal contaminants to accelerate the occurrence of pitting and spalling; the use of the relevant sensors on the bearing movement vibration signals, temperature and other state data synchronized collection; the data signal decomposition process, based on the decomposition results to extract bearing wear characteristics, build wear feature set; for bearing wear to build wear level, build a wear monitoring model will be input into the model, the model training results and different levels of threshold comparison, to determine the severity of bearing wear, complete the bearing wear monitoring. Wear level is constructed for bearing wear, a wear monitoring model is constructed to input the features into the model, and the training results of the model are compared with the thresholds set for different levels to determine the severity of bearing wear and complete the monitoring of bearing wear. Due to the large number of noisy data signals in the process of collecting state data signals, the detection effect of this method in bearing wear monitoring is not satisfactory.

Based on the above mentioned drawbacks in the bearing raceway wear fault monitoring process, an algorithm is proposed to monitor the large-area deep wear faults of the yaw bearing raceway of wind turbines. Vibration data are collected through accelerometers. After missing value interpolation and improved wavelet mode maximum denoising, the subdomain adaptive adversarial transfer network is used to extract fault features and classify and identify them, and a wear grade scoring and early warning mechanism is established. Since traditional methods assume that the noise follows a Gaussian distribution and is stationary, while the impact noise in the yaw bearing vibration signal has time-varying amplitudes and suddenness (such as the wideband transient response generated at the moment of raceway spalling). Fixed-threshold denoising may mistakenly filter out the fault impact components as noise, resulting in the loss of fault features. The improved modular maximum denoising algorithm can adaptively adjust the thresholds of each scale based on the local variance of the noise, avoid excessive suppression of the impact components, and effectively distinguish the noise (randomly occurring) from the fault characteristics (cross-scale correlated).

2. Wind turbine yaw bearing state data signal acquisition and pre-processing

The yaw bearing is one of the key components of the wind turbine, responsible for rotating the nacelle of the wind turbine to stand against the wind to achieve the best power generation efficiency. Wear of the yaw bearing raceway will directly affect the stability and safety of wind turbine operation [9-10]. By monitoring the wear of the yaw bearing races, potential failures can be detected and warned in time, so that repair or replacement measures can be taken in advance, avoiding the loss of power generation efficiency and the increase of maintenance costs caused by the failure and shutdown. At the same time, regular monitoring and

maintenance can also extend the service life of the wind turbine, improve its economic performance; at the same time, the monitoring of bearing raceway wear, but also for the repair and replacement work to provide an accurate basis for the analysis of the monitoring data to determine the degree of wear and failure of the bearings, so as to formulate targeted maintenance programs. This not only improves the maintenance efficiency, but also ensures the quality of maintenance and reduces the risk of equipment damage caused by improper maintenance.

Based on the above analysis results, the yaw bearing raceway large-area deep wear fault monitoring of wind turbine is divided into three parts: bearing state data acquisition, raceway wear fault feature extraction and fault monitoring and warning. Among them, the bearing running process state data acquisition is the key link of the whole monitoring method, and the specific design is as follows:

2.1. Bearing-related state data acquisition

When monitoring the failure of yaw bearing raceway in large area and deep wear, it is necessary to select suitable sensors to collect the data of bearing operation process. As the bearing wears, the bearing wear will generate specific vibration frequencies, which may include the fundamental frequency of the bearing, harmonics and sidebands, etc. The vibration sensor can capture these frequency components to analyze the bearing's operating condition.

Firstly, based on the WTG bearings in operation, vibration sensors and temperature sensors with permanent or portable magnetic bases are mounted on the bearing box for data acquisition [11-12].

Wind turbine yaw vibration signal acquisition mainly use piezoelectric accelerometer (vibration sensor) to expand and during the process of bearing work, the charge Q generated by the mechanical stress of the piezoelectric material for setting the accelerometer is proportional to the stress K_p. The generation process of the charge is as follows:

$$Q = K_p \cdot a \tag{1}$$

Where, the acceleration applied to the sensor itself is expressed as a form of a.

After charge acquisition is completed, a charge amplifier is used and its amplification gain is set to G, by which charge is converted to voltage V_{out}, the process is shown in the following equation:

$$V_{out} = G \cdot Q = G \cdot K_p \cdot a \tag{2}$$

Upon completion, signal conditioning is applied to the voltage to obtain the final bearing vibration data signal (electrical signal) V_{final}. The results are shown in the following equation:

$$V_{final} = H(V_{out}) = H\left(G \cdot K_p \cdot a\right) \tag{3}$$

Where the signal transfer function is expressed as a form of H.

2.2. Missing value interpolation

Due to the vibration sensor to collect the bearing state data information, based on the sensor's own interference, the data collected will be partially missing, so after the collection of data, you need to use the appropriate algorithms for the data collected to implement the missing value interpolation processing.

The definition of data missingness varies in different fields, and for bearing state data, in order to avoid confusion, the sparsity of the dataset is used to describe the data missingness. Sparsity refers to the ratio of the number of non-missing units in a data set to the total number of units. The sparser the data set, the smaller the sparsity value, the corresponding missing degree value is larger, generally considered the sparsity of 0.5 or less is called a sparse data set.

2.2.1. Data clustering

Due to the existence of missing values and the incomplete data structure, directly filling in the missing values will result in poor filling effects due to the complexity and uncertainty of the data distribution. The fuzzy C-means clustering algorithm is adopted to cluster the collected state data, which can group the data with similar features into one category, making the data within the same category closer in terms of distribution and features. When filling in the data subsequently, by taking advantage of the characteristics within the data clusters after clustering, appropriate reference attribute variables are selected based on correlation to fill in the missing values. Within the data clusters after clustering, the reference attribute variables with the highest correlation to the attribute variables to be filled can be found more accurately. Because clustering makes the data have a stronger correlation within a local range, the selection of reference attribute variables is more reasonable. Thus, the average ratio can be calculated more

accurately based on the local characteristics of the data after clustering, achieving more effective filling of missing values. Therefore, clustering is performed before imputation. By rationally grouping the data, clustering provides a more reliable data basis for the subsequent imputation based on the average ratio method of correlation, assists and influences the subsequent imputation process, and effectively improves the accuracy and rationality of missing value imputation.

Firstly, the fuzzy C-mean clustering algorithm is used to implement the clustering process for the collected state data. The process is as follows:

Step 1: Based on the collected m bearing status data, first determined c number of clusters, initialize the affiliation matrix $U^{(r)}$, and make it satisfy the condition of the following equation, labeling the iterative process as $r = 0,1,\cdots$, as shown in the following equation:

$$\begin{cases} \sum_{i=1}^{c} u_{ij} = 1; \forall j = 1,\cdots,m \\ \sum_{i=1}^{c} u_{ij} > 0; \forall i = 1,\cdots,c \end{cases} \quad (4)$$

Where, the membership of the i th sample x_i belonging to the central point of the j th cluster c_i is expressed with u_{ij}.

Step 2: Based on the affiliation matrix $U^{(r)}$, calculate the center $c(k,d)$ of clustering of the data, the process is shown in the following equation:

$$c(k,d) = \frac{\sum_{i=1}^{m} u_{ik}^{n} x_{i,d}}{\sum_{i=1}^{m} u_{ik}^{n}} \quad (5)$$

Where, the membership of the sample x_i to the cluster center c_k is expressed as $u_{i,k}$, and the d dimensional sitting mark of the corresponding sample x_i is made as $x_{i,d}$, the total number of samples is denoted as m, the fuzzy factor is denoted as n.

Step 3: According to the current clustering center, update the affiliation matrix, the process is shown in the following equation:

$$U_{ik} = 1/\sum_{j=1}^{c} \left(\frac{\|x_i - C_k\|}{\|x_i - C_j\|} \right)^{\frac{2}{n-1}} \quad (6)$$

Where, the representation vector of the i th data is denoted as x_i, the k th cluster center is denoted as C_k, the j th cluster center is denoted as C_j, and the Euclidean distance between the data and the cluster center is denoted as $\|x_i - C_k\|$, the fuzzy factor is denoted as n.

Step 4: Repeat the above steps 2 and 3 until the stopping criterion is satisfied and the data clustering is realized.

2.2.2. Populated data

After completing the clustering of the data, the data were populated using the average ratio method, as shown in the following equation:

Step 1: Firstly, the correlation calculation is used to obtain the correlation of the attribute variables of the bearing state dataset within the clustered data clusters;

Step 2: Select attribute variables that contain one or more missing values: Z_1 (attribute variable to be populated).

Step 3:Select the reference attribute variables (additional variables), which satisfy the highest correlation with Z_1, and contain complete values or at least does not have the same missing value relative to Z_1, denoted as Z_2;

Step 3: Calculate the average ratio ARV of the variables. the process is shown in the following equation:

$$ARV = 1/n \sum_{i=1}^{n} (Z_{i2}, Z_{i1}) \quad (7)$$

In the formula,$n$indicates the number of samples in the data cluster that do not contain missing values, Z_{i1}, Z_{i2} denotes the i th taken value of the attribute variable Z_1, Z_2;

Step 5: Calculate the padding value, during the process, if the ratio is calculated by (Z_1/Z_2), then the missing value of the attribute Z_1 is computed as $Z_2 * ARV$, if the ratio is calculated by (Z_2/Z_1), then the missing value of the attribute Z_1 is computed as Z_2/ARV;

Step 6: Repeat the above steps until all the missing values of each attribute variable are filled in, i.e. the missing value interpolation of the bearing state data is completed.

For the attribute variable Z_1 containing multiple missing values, the method of sequential filling is adopted. Each time when filling, according to the process of steps 1 to 5, the average ratio ARV is calculated using the reference attribute variable Z_2 with the highest correlation to Z_1 and no identical missing values, and then the missing value filling value of the Z_1 attribute is calculated. After completing the filling of one missing value, update the data set and repeat the above steps for the next missing value until all the missing values in Z_1 have been filled. Since the filling process will change the overall distribution of the data set and the correlation between variables, each filling will be affected by the previously filled values. In subsequent steps such as the selection of reference attribute variables and the calculation of the average ratio, the impact brought by this change will be reflected, thereby

making the filling process more in line with the actual situation and coherent. This clearly expounds the mutual influence relationship among multiple missing values.

2.3. Pre-processing of bearing state data signal

Due to the use of sensors to collect wind turbine yaw bearing vibration and temperature data, the sensor itself there is a certain amount of interference, the collection of data signals will be mixed with a large number of noise signals, so it is necessary to implement the de-noising process to improve the signal quality of the collected data signals.

During the process, CEEMDAN algorithm is used to decompose the collected bearing vibration data signals, and the inherent modal components of the decomposed signals are obtained to carry out the improved wavelet maxima denoising, which approximates the useful signals and realizes the quality enhancement of the bearing state data signals [13-14].

2.3.1. Signal decomposition of bearing state data

The EMD algorithm can decompose any signal into a number of intrinsic modal components (IMF) and a residual residue, the IMF frequency from high to low, and the residual residue is monotonous. Gaussian white noise has a uniform power spectral density. Its addition can cause differentiated disturbances in the similar frequency components of the original signal. According to the extreme point screening characteristics of EMD, this disturbance can effectively separate the originally overlapping instantaneous frequency components (that is, solve the modal aliasing). Through multiple averaging processing of positive and negative noise pairs, the IMF components of the noise itself will cancel each other out, while the IMF components of the true signal are retained, thereby significantly reducing the residual noise and improving the decomposition accuracy. For this reason,the CEEMDAN algorithm adds the positive and negative Gaussian white noise of each order of the IMF components decomposed by the EMD in the decomposition process, which can eliminate the problem of modal overlapping to a large extent, improve the decomposition efficiency, and reduce the noise residual and reconstruction error. residual noise and reconstruction error.

CEEMDAN is broken down in the following steps:

Step 1: Add K Gaussian white noise signal to the original signal, and the state data signal $x^i(t)$ after noise addition is as follows:

$$x^i(t) = x(t) + (-1)^m \omega_0 \varepsilon^i(t); i \in 1,2,\cdots,K; m \in \{1,2\} \quad (8)$$

Where, the noise factor is expressed as a form of ω_0, independent positive and negative paired Gaussian noise signals with variance 1 are formulated as $(-1)^m \omega_0 \varepsilon^i(t)$, the number of independent positive and negative pairs of Gaussian noise is denoted by m.

Step 2: For each $x^i(t)$, which is decomposed using the EMD decomposition algorithm to obtain the first of the data signals $IMF1$, and set the total number of $IMF1$ is K, for K individual $IMF1$ mean value expand calculation $c_1(t)$, and by calculating the difference between the result and the original data signal $x(t)$, the first residual signal $r_1(t)$ is obtained, as follows:

$$\begin{cases} c_1(t) = \frac{1}{K}\sum_{i=1}^{K} c_1^i \\ r_1(t) = x(t) - c_1(t) \end{cases} \quad (9)$$

Step 3: Decompose the residual signal $r_1(t) + \omega_1 E_1(\varepsilon^i(t))$ of the added noise, where, E_j represents the jth IMF2 component obtained from the EMD decomposition result after adding Gaussian white noise. After decomposition, there are still a number of K individual $IMF2$,and then expand the average $c_2(t)$ of these K individual $IMF2$ to calculate. The results are as follows:

$$c_2(t) = \frac{1}{K}\sum_{i=1}^{K} E_1(r_1(t) + \omega_1 E_1(\varepsilon^i(t))) \quad (10)$$

Step 4: Repeat step 3 above until the residual fraction $r_n(t)$ is not possible to proceed to the next step of disaggregation, i.e. The number of extreme points $r_n(t)$ is less than 2; at this point, set acquires a number of K individual IMF component $c_K(t)$, with the remaining residuals set to $R(t)$, i.e., the original signal $s(t)$ is expressed in the following equation:

$$s(t) = \sum_{K=1}^{K} c_K(t) + R(t) \quad (11)$$

2.3.2. Wavelet maxima denoising

(1) The principle of wavelet mode maxima denoising

The Lie index is a mathematical measure that characterizes the local features of a function. The Lee's exponent of a function at a point characterizes the size of the singularity at that point, the larger the exponent, the higher the smoothness of the point, the smaller the exponent, the larger the singularity at that point. Usually, the Lee's index of impulse

noise and white noise appearing in one-dimensional signals are less than -1/2, and the singularity is larger. When the index δ is greater than 0, the extreme value of the wavelet transform increase with the increase of decomposition scales j; when δ is less than 0, the extreme value of the wavelet transform decreases with the increase of the decomposition scales j. It can be seen that the mode maximum of the wavelet coefficient obtained by the signal point increases with the increase of the decomposition scale j, while the mode maximum of the wavelet coefficient obtained by the noise point decreases with the increase of the decomposition scale j, which should be filtered out.

Bearing vibration signal and noise signal in different scales under the wavelet changes show different characteristics, can use this different characteristics to set a certain threshold to filter out the noise signal, and then the remaining wavelet decomposition of the signal reconstruction to realize the signal denoising.

In terms of the threshold estimation method, an adaptive threshold strategy based on the statistical characteristics of noise is adopted. Statistical analysis was conducted on the noise wavelet coefficients of each scale j. Assuming that the noise follows a generalized Gaussian distribution, according to the general threshold criterion, a scale correlation correction factor is introduced to dynamically adjust the initial threshold. This multi-scale adaptive mechanism can effectively distinguish the impact components of bearing faults (manifested as the modulus maxima correlated across scales) from random noise (occurring only at a single scale). The specific calculation process is as follows:

$$T_j = \alpha_j \cdot \sigma_j \cdot \sqrt{2\,\boldsymbol{ln}\,N} \cdot s(t) \tag{12}$$

In the formula, N represents the signal length, α_j represents the scale correlation correction factor, σ_j represents the median estimate of the wavelet coefficient of the j layer.

(2) An improved wavelet mode-maximization denoising algorithm

Accurately obtaining the wavelet mode maxima requires constructing a domain around the mode maxima to perform the mode maxima tracking. Therefore, constructing an effective field is an important factor affecting the tracking speed. In order to improve the efficiency of tracking the wavelet mode maxima of useful signals, wavelet shrinkage is utilized to remove the mode maxima corresponding to certain noises, and then the mode maxima points at each level are tracked according to the wavelet mode maxima propagation of the signals. The wavelet thresholds for each level are estimated based on the wavelet coefficient distributions of the signal and noise at different levels. The threshold value is used to obtain the wavelet transform mode maxima of the signal, reduce the number of fields to be tracked, and control the effective field range appropriately, so as to improve the tracking efficiency.

The selection of the threshold function, the traditional hard threshold function $d_{j,k}^{y,\wedge}$, the soft threshold function $d_{j,k}^{r,\wedge}$ and the Garrote function $d_{j,k}^{G,\wedge}$ are generally selected, as shown in the following equation:

$$\begin{cases} d_{j,k}^{y,\wedge} = \begin{cases} d_{j,k}; |d_{j,k}| \geq \lambda \\ 0; |d_{j,k}| < \lambda \end{cases} \\ d_{j,k}^{r,\wedge} = \begin{cases} d_{j,k} - \lambda \mathrm{sgn}(d_{j,k}); |d_{j,k}| \geq \lambda \\ 0; |d_{j,k}| < \lambda \end{cases} \\ d_{j,k}^{G,\wedge} = \begin{cases} d_{j,k} - \frac{\lambda^2}{d_{j,k}}; |d_{j,k}| \geq \lambda \\ 0; |d_{j,k}| < \lambda \end{cases} \end{cases} \tag{13}$$

Where, the signal wavelet estimation coefficients is described as a form of $d_{j,k}^{\wedge}$, the original wavelet coefficients is described as a form of $d_{j,k}$, the threshold is chosen to be described as a form of λ, the wavelet decomposition scale is denoted by j, the integer coefficients is denoted as k.

When using the above threshold function for denoising, due to the hard threshold function to process the signal, the edge of the denoised signal is better, and can well retain the singularity of the useful signal, but due to the discontinuity of the function itself, it may lead to a large oscillation in the processed signal, resulting in a pseudo-Gibbs effect. The soft threshold function has a relatively good continuity of the processed signal, the denoising effect is smoother, but the overall wavelet coefficients and the original coefficients of the overall deviation of the original coefficients, which will make the signal of the high-frequency part of the signal is lost, resulting in the reconstruction of the signal distortion. The Garrote function combines the advantages of soft and hard threshold functions, which can not only maintain the continuity of the signal, but also retain the high-frequency information of the signal, overcoming the shortcomings of the soft and hard threshold functions. However, for the mutation

signal, its denoising effect is still insufficient.

In order to solve the disadvantages of the above three threshold functions, an exponential threshold function is given. Considering the high order derivability of the natural exponential function, the natural exponent can be introduced into the threshold function to overcome the disadvantages of the hard threshold function. The exponential threshold function is shown in the following equation:

$$d_{j,k}^{ZS,\wedge} = \begin{cases} d_{j,k} - \frac{\lambda^2}{d_{j,k}^{Q}}; \left|d_{j,k}\right| \geq \lambda \\ 0; \left|d_{j,k}\right| < \lambda \end{cases} \quad (14)$$

Where, the adjustment is described as a form of Q.

In the denoising process, because the mode maximum amplitude of the noise decreases at a binary rate as the scale increases, which results in the mode maximum points on the largest scale being controlled by the signal, while some smaller amplitudes may still be propagated by the noise maximum points on the lower scale, for this reason, the threshold T is set for the maximum point on the maximum scale 2^J, the module maximum point below the threshold T is removed, and the point noise mode maximum is dominant. The threshold T is specifically in the $T = O\frac{\max\left|d_{2^J}f(x_i)\right|}{J}$ form, where the J layer wavelet coefficient of the signal is described in the $d_{2^J}f(x_i)$ form and the constant coefficient in the O form.

According to the threshold value determined above as well as the exponential threshold function, the noise of the signal is effectively filtered out, and signal denoising is realized through signal reconstruction to improve the signal quality [15-16]. The specific process is shown in Fig. 1.

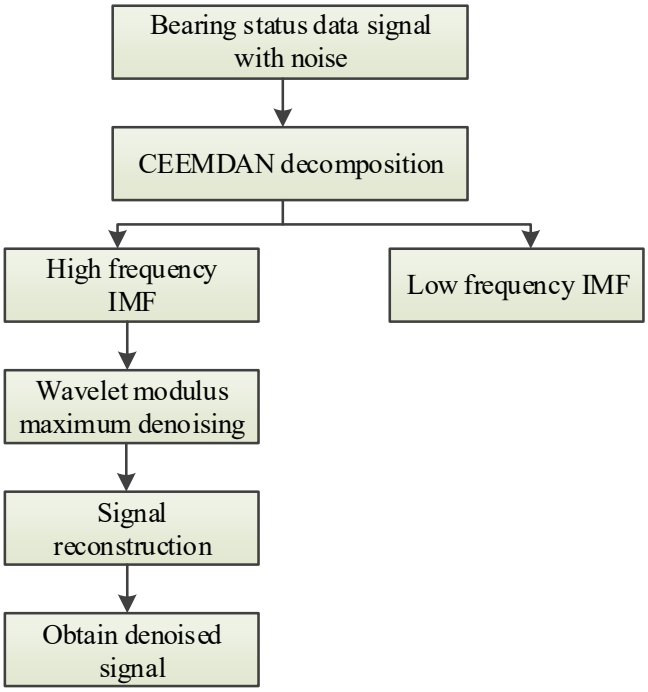

Figure 1. Flow chart of denoising bearing state data signal.

3. Yaw bearing raceway extensive wear failure monitoring algorithm design

For the collected vibration data signal after denoising, the bearing wear fault characteristics are extracted, and a diagnostic model is constructed to classify and recognize the extracted characteristics [17-18] at the same time, the wear severity level is set according to the bearing wear level, and an early warning mechanism is established to score the wear level of the classified faults, find out the corresponding wear level, and identify whether early warning is needed; finally, when the bearing raceway damage exceeds the early warning threshold, the early warning mechanism is activated to notify the maintenance personnel to replace the bearing for repair treatment, so as to realize the real-time monitoring of the large-scale wear faults of the yaw bearing raceway of wind turbine generators.

3.1. Classification algorithm for bearing fault feature extraction

After determining the bearing state, the vibration data signals collected from the unhealthy bearings are used to extract the fault characteristics of the bearings through the improved domain adversarial network.

A bearing fault feature extraction classification algorithm model based on a subdomain-adapted adversarial migration

network [19-20] is proposed for bearing raceway wear monitoring. Firstly, a basic 1DCNN model is trained in pre-training with sufficient source domain data to complete the depth mapping from the original vibration signals to the fault features; subsequently, an improved domain adversarial model is built in the adversarial training stage to learn the domain-invariant features; The Wasserslein distance is introduced in the adversarial layer to measure the distance between the extracted features in the source and target domains, and to utilize the advantage of Wasserslein distance in the gradient to obtain stable training results. The local maximum mean deviation is introduced in the classification layer to measure the difference between the source and target domain data embedded in the relevant subdomains, in order to capture the fine-grained information of each category and adjust the distribution of the relevant subdomains under the same category. The domain-invariant features are learned in constant confrontation by maximizing the domain discrimination loss and minimizing the loss of the classifier and feature extractor. The adversarial learning process enables end-to-end unsupervised bearing fault migration diagnosis by introducing labeled source-domain migratable features into new and relevant target-domain diagnostic tasks without any labeled samples.

The specific bearing fault feature extraction classification model is shown in Fig. 2.

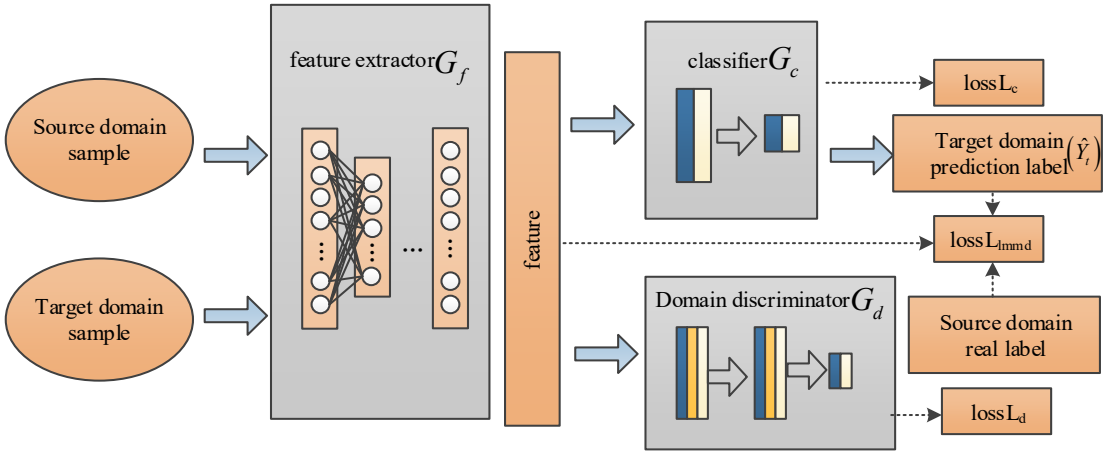

Figure 2. Classification model for bearing fault feature extraction based on subdomain adaptive adversarial transfer network.

Analyzing Fig. 2, it can be seen that the model is divided into three parts, including feature extractor, domain discriminator and classifier. The feature extractor is constructed by the basic CNN model, while the domain discriminator and classifier are constructed by 3-layer and 2-layer shallow fully connected networks, respectively. In the feature extraction part, the 1DCNN model is used to automatically extract high-dimensional deep features from the preprocessed signals in both domains, and its parameters are shared and updated by the classifier, domain discriminator and the introduced LMMD computational module at the same time. In the domain discriminator part, WD is introduced to maximize the discriminator error. In the classification part, the cross moisture loss is used to minimize the source domain classifier error. The training process needs to be based on pre-training, using the source domain data to train the feature extractors and classifiers to obtain the pre-training parameters, and loading the pre-training parameters into the SANN model to complete the initialization of the model parameters during the formal training. The specific process is as follows:

3.2.1. Feature extraction

In the feature extraction process, the above denoised bearing state data are input into the model, and the feature extraction module is designed as a 4-layer one-dimensional convolutional neural network, using the ReLU function as the activation function. For the one-dimensional vibration signals, in order to obtain a rich sensory field and better extract the global features of the signals, a large-inch convolution kernel is used for the 1st convolutional layer; a maximal pooling layer and adaptive pooling layer are added to the 2nd and 4th convolutional layers, respectively, for sampling operations to reduce the size of the parameter matrix, in which the output dimension of the adaptive pooling layer is 4.

In order to enhance the generalization ability of the model and reduce the model complexity, a batch layer and a Dropot

layer are introduced, and a BN layer is added after each convolutional layer respectively, which effectively avoids the problem of gradient disappearance and improves the training speed of the model. The Dropot of the fully connected layer is set to 0.5, which effectively improves the generalization ability of the model. The specific parameter settings are shown in Table 1.

Table 1. Network Structure of DCNN.

Layer type	Nuclear size and quantity	Step
Convolutional layer1	16*15	1
Convolutional layer2	32*2	1
Maximum pooling layer	2	2
Convolutional layer3	64*3	1
Convolutional layer4	128*3	1
Adaptive pooling layer	—	—
Fully connected layer	256	—

Among them, the typical manifestation of raceway wear in yaw bearings is periodic impact signals, and the pulse width is positively correlated with the damage size. At a sampling rate of 20kHz, the impact pulse width of slight wear is approximately 15 to 20 sampling points. The first layer adopts a large convolution kernel design of 16*15. The width of a single convolution kernel can completely cover the time-domain waveform of an impact event, avoiding feature truncation. Raceway wear can stimulate characteristic frequencies and their harmonic components, and different degrees of damage can lead to a redistribution of frequency energy. In the shallow network (Conv1-2), the small-step long volume volume of 16-32 channels focuses on the extraction of single-cycle impact patterns, corresponding to the capture of high-frequency components (2-8 KHZ). The maximum pooling layer (step size 2) enhances the significance of the impact amplitude by retaining local extremum values; In the deep network (Conv3-4), the convolutional layers of 64-128 channels expand through receptive fields (up to 128 points), learn multi-period correlation patterns, and effectively capture low-frequency modulation features (<1kHz). The output dimension of the adaptive pooling layer is set to 4, corresponding to the classification of the degree of wear, forcing the network to establish a mapping relationship between the frequency band energy and the degree of failure. Considering that the traditional average pooling may blur the impact characteristics, this design adopts a hybrid pooling strategy. The peak amplitude of the impact is retained in the maximum pooling layer. In the adaptive pooling layer: the downsampling strategy is dynamically adjusted through learnable weights to avoid the frequency shift caused by rotational speed fluctuations. Due to the interferences such as rotational speed fluctuations and load changes in the on-site data, batch normalization is adopted to make the network focus on the relative energy distribution rather than the absolute amplitude, reducing the fluctuation of classification accuracy under variable rotational speed conditions. And by randomly shielding neurons, the network is forced to rely on both the time-domain impact characteristics and the frequency-domain harmonic distribution simultaneously, avoiding overfitting to a single mode.

The convolutional layer is the core of CNN, which is used to extract the high dimensional features of the input, and consists of multiple convolutional kernels, when each kernel slides on the input mapping, it shares the weights and biases, and obtains the feature map by convolving the input signal with the convolutional kernel. After the convolution operation, the activation function is used to increase the nonlinear representation of the network. The output of the convolutional layer is:

$$C_n^l = f(\sum_{i=1}^{m} x_n^{l-1} * \omega_n^l + b_n^l) f(x) = max(0, x) \qquad (15)$$

Where, the ReLU activation function is expressed in the form $f(x)$, the i th feature graph output by the module $l-1$ layer is expressed in the form x_n^{l-1}, the n convolution of the l th layer is described as ω_n^l, the shared bias is expressed in the b_n^l form, and the nonlinear output of the n feature graph of layer 1 is described in the C_n^l form after the activation function mapping.

In addition, the main role of the pooling layer is feature extraction. By removing the unimportant samples in the feature map, the number of parameters is further reduced, and overfitting is also controlled to a certain extent. Common pooling layers include maximum pooling and average pooling.

As for the fully connected layer, the first part of the convolution and pooling is equivalent to the feature engineering, and the latter part of the fully connected layer is equivalent to the feature weighting, which plays the role of classifier in the whole convolutional neural network. Finally, after the Soflmax layer, the probability distribution of the

current sample belonging to different categories can be obtained.

3.2.2. Adversarial training based on domain discriminators

In this stage, Wasserslein distance can be introduced based on the network adversarial layer to carry out effective distance metrics on the features of the extracted source domain data and target domain data, and using the metric results, feature edge alignment can be realized, and the LMMD computing module can be introduced to obtain the fine-grained information of the feature categories to realize the conditional distribution alignment.

(1) The Wasserslein distance metric

The collected bearing status data is set into two sample types, source domain X^s and target domain X^s, and the two samples are passed through the network parameter ϕ_f feature extractor r_f , mapped into high-dimensional features, to generate source-domain features $h_s = r_f(x_s)$ and the target domain characteristics $h_t = r_f(x_t)$. And order the distribution of the edges of h_s and h_t as P_s and P_t, the result of feature mapping for the source domain as well as the target domain for the setting domain discriminator r_d are $r_d(h_s)$ and $r_d(h_t)$, and thus the calculations are carried out to the difference $W(P_s, P_t)$ in feature distribution between the source and target domains by means of the WD algorithm as shown in the following equation:

$$W(P_s, P_t) = \sup_{\|r_d\| \le 1} E_{h_s \sim P_s}[r_d(h_s)] - E_{h_t \sim P_t}[r_d(h_t)] \quad (16)$$

Where, the upper boundary is described as sup form, the role of "sup" is that it is effective for all domain discriminators that satisfy the conditions of the 1-Lipschitz continuous function. It is an operation used when calculating the difference in the feature distribution between the source domain and the target domain. In simple terms, it is to find an upper bound among the calculation results of numerous domain discriminators that meet specific conditions. This upper bound is involved in the quantification process of the difference in feature distribution between the source domain and the target domain, thereby helping to analyze the difference in feature distribution between the source domain and the target domain.The expectation of the value obtained by the r_d from the feature h_s randomly selected from the source domain feature distribution P_s is expressed in $E_{h_s \sim P_s}[r_d(h_s)]$, and the expectation of the value obtained by the r_d from the feature h_t randomly selected from the target domain feature distribution P_t is expressed in $E_{h_t \sim P_t}[r_d(h_t)]$. The sup is valid for all those satisfied with the 1-Lipschitz continuous function r_d, at this point, the empirical approximation L_{wd} of the WD is estimated as follows:

$$L_{wd} = \frac{1}{n_s}\sum_{x_s \in X_s} r_d\left(r_f(x_s)\right) - \frac{1}{n_t}\sum_{x_t \in X_t} r_d\left(r_f(x_t)\right) \quad (17)$$

During the training process, the parameter of r_d are continuously updated to ensure that the WD of the features in the source and target domains are maximized, whereas, the direction of parameter r_f optimization is to minimize the WD of the features in the source and target domains to achieve the purpose of edge distribution adaptation.

To solve the maximization optimization problem when satisfying the 1-Lipschitz continuum constraint. A spectral normalization method is used to solve the problem of discriminator oscillations during adversarial training, i.e., the BN method is normalized with spectral paradigms instead of the BN method at each layer of the discriminator network. When using spectral normalization in the discriminator network, the singular value decomposition is performed on the weight matrix of each layer of the network to obtain its maximum singular value as the spectral norm. Normalize the weight matrix to ensure that each layer mapping satisfies the constraint condition that the Lipschitz constant does not exceed 1; The spectral norm is efficiently estimated through the power iteration method. This method only requires 3 to 5 iterations to converge, significantly reducing the computational overhead. Compared with batch normalization (BN), spectral normalization has two major theoretical advantages: First, it fundamentally avoids the gradient explosion problem by explicitly controlling the Lipschitz constant, while BN only performs implicit adjustment through empirical statistics; Secondly, maintain the stability of the discriminator's decision boundary. Specifically in the network implementation, a spectral normalization layer is inserted after each convolutional layer and fully connected layer of the discriminator, while the original BN layer and Dropout layer are removed. This design enables the discriminator to continuously provide a stable gradient signal during the

adversarial training process.

(2) LMMD calculations

The main focus of the adversarial domain adaptation network is on edge distribution alignment, while ignoring the relationship between two subdomains in the same category, in fact, the data of the same fault category have stronger correlation. When only focusing on edge alignment for adaptation, the subdomain features are too close to each other, while after conditional distribution adaptation, both global and subdomain features achieve distributional alignment.

LMMD is introduced to solve the fine-grained problem under global alignment and realize the alignment of the same class of feature distribution, the alignment process is shown in the following equation:

$$\begin{cases} L_c = \frac{1}{C}\sum_{c=1}^{C}\left\|\sum_{x_i^s\in D_s} w_i^{sc}\varphi(x_i^s) - \sum_{x_i^s\in D_s} w_j^{sc}\varphi(x_j^t)\right\|_H^2 \\ w_i^c = \frac{y_{ic}}{\sum_{(x_j,y_j)} y_{jc}} \end{cases} \quad (18)$$

Where, the corresponding weights of x_i^s, x_j^t belonging to category c are described in the form of w_i^{sc}, w_j^{sc}, the number of fault categories is described by c and the feature alignment coefficients in φ.

In setting the output of the sample feature extractor in the source and target domains as z_{si} and z_{ti}, using the kernel function to map the two to regenerative and Hilbert spaces, and to compute the local maximum mean error L_c of the sample features in the source and target domains. The results are shown in the following equation.

$$L_c = \frac{1}{C}\sum_{c=1}^{C}\left[\sum_{i,j=1}^{n} w_i^{SC} w_j^{SC} k(z_{si}, z_{sj}) + \sum_{i,j=1}^{n} w_i^{tc} w_j^{tc} k(z_{ti}, z_{tj}) - 2\sum_{i,j=1}^{n} w_i^{sc} w_j^{tc} k(z_{si}, z_{tj})\right] \quad (19)$$

3.2.3. Classifiers

Finally, the aligned features are fed into the classifier. Based on the loss function L of the classifier, the classification of characteristics is completed to determine the wear category of the bearing raceway. The specific classification process is shown in the following equation:

$$\begin{cases} y_{i,c} = L(\lg(p_{i,c}))x_{i,j} \\ L = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c}\lg(p_{i,c}) \end{cases} \quad (20)$$

In the formula, the total number of collected bearing condition data is denoted by N, total number of wear categories is denoted by C, the classifier prediction probability is formulated as $p_{i,c}$, the sample prediction results are described as a form of $y_{i,c}$.

3.3. Designing the early warning mechanism for yaw bearing wear in wind turbines

Aiming at the type of yaw bearing failure, the different degree of bearing wear is combined to formulate a wear severity level of wind turbine yaw bearing raceway, and the results are shown in Table 2.

Table 2. Wear grades of yaw bearing raceway for wind turbines.

Wear grade	State	Score	Describe
1	Mild wear	[0-0.3]	Small area wear, not exceeding 10% of the total area of the raceway, with a depth less than 10 μ m
2	Moderate wear and tear	[0.3-0.5]	Large area, accounting for 10% -30% of the total raceway area, with a depth between 10 μ m and 50 μ m
3	Severe wear and tear	[0.5-0.8]	Wide area, accounting for 30% -50% of the total raceway area, with a depth ranging from 50 μ m to 100 μ m
4	Disaster level wear and tear	[0.8-1]	The surface of the Extensive wear, exceeding 50% of the total area of the raceway and a depth of over 100 μ m.

Based on the bearing raceway wear level designed in Table 2, the level 2 moderate wear is set as the warning value, when the wear characteristics are determined, the fraction calculation is carried out, when the result $\rho \geq [0.3 - 0.5]$, it indicates that the bearing races are worn over a large area and to a high depth, so it is necessary to issue an early warning and carry out replacement repairs.

When unfolding the scoring computation for the classified wear features, it is assumed that extract n fault characteristics, denoted as vector $Y = \{y_1, y_2, \cdots, y_n\}$, with its corresponding labeling q of the degree of bearing wear (This label may have been obtained through expert scoring, historical data, or other means). Our goal is to find a mapping function that will map Y to a score S of Y. The process of the specific wear characterization score S is shown in the following equation:

$$S = W_i * y_i + b \quad (21)$$

Where, the weight of the i th wear fault feature is expressed in a form of W_i, and the bias term is described in b.

Finally, according to the calculated score, the wear level of the corresponding bearing raceway is found, realizing the real-time monitoring and warning of the large-area wear of the bearing raceway.

4. Experiment

In order to verify the overall effectiveness of the above bearing raceway wear monitoring method, it is necessary to test the method.

The wind turbine yaw bearing raceway large-area deep wear failure monitoring algorithm (proposed method), A new method for diagnosing faults in the outer raceways of rolling bearings in asynchronous motors, Bearing fault diagnosis of pumped storage units considering combined acoustic-vibration modes, A quantitative bearing fault diagnosis method based on MTF and improved residual network, Fault diagnosis for abnormal wear of rolling element bearing fusing oil debris monitoring are selected for comparative tests, to verify the feasibility of practical application of the proposed method in the monitoring of large-area wear of bearing raceways.

During the experiment, the wind turbine of a city wind power station is selected as the test target, and a piezoelectric accelerometer and a temperature sensor are used to collect the state data of the yaw bearing of the wind turbine and integrate the collected data to construct a test sample data set. Among them, piezoelectric sensors have a wider effective frequency band and can better capture the low-frequency flutter and high-frequency impact signals unique to yaw bearings. Their high sensitivity and anti-electromagnetic interference characteristics make them particularly suitable for the complex electromagnetic environment inside wind turbine towers. In contrast, although piezoresistive sensors perform stably in the low-frequency range, their limited dynamic range and relatively low resonant frequency make it difficult to accurately characterize the high-frequency vibration characteristics caused by bearing raceway wear, and they are also susceptible to temperature drift. These limitations make them less applicable in scenarios that require wide-band monitoring, such as yaw bearings. During the test, the platform equipped with Ubuntu 20.04 LTS operating system to simulate the actual wind turbine operation state is used as the experimental environment, and the specific experimental parameters are set as follows:

Bearing operating parameters:

(1) Contact load: 10kN.

(2) Swing angle: 0.8°8, 1.2°1, 1.5°4.

(3) Oscillation frequency: 5Hz.

(4) Experimental temperature: about 20℃.

(5) Relative humidity: about 50%.

(6) Number of experimental cycles: 10^3 times.

AKG392 Piezoelectric Acceleration Sensor

(1)Operating temperature:-196℃~+200℃;

(2)Measuring range:0.001~800MPa.

(3) Sensitivity: 0.21000PC/MPa.

(4) Intrinsic frequency: 75~500kHz.

Improvement of domain adversarial network parameterization:

(1) Learning rate: 0.001;

(2) Batch size: 128;

(3) Number of training rounds: 50;

After the parameters were set, based on the experimental parameters set above, five sets of data were selected from the test sample data set to test the detection effectiveness of the large-area deep wear monitoring method for bearing tracks.

4.1. Test of denoising effect of bearing state data

When carrying out the monitoring of large-area deep wear of bearing raceway, the bearing state data collection as the core foundation of wear monitoring, if the quality of the collected data is poor, it will directly affect the actual accuracy of the subsequent monitoring of bearing raceway wear. For this reason, A new method for diagnosing faults in the outer raceways of rolling bearings in asynchronous motorsproposed proposed in literature [5] is set up as the comparison method 1, Bearing fault diagnosis of pumped storage units considering combined acoustic-vibration modesmethod proposed in literature [6] is set up as the comparison method 2, A quantitative bearing fault diagnosis method based on MTF and improved residual networkproposed in literature [7] is set up as comparison method 3, Fault diagnosis for abnormal wear of rolling element bearing fusing oil debris monitoring

proposed in literature [8] is set up as the comparison method 4, joint all these proposed method to carry out the bearing raceway large area wear monitoring, any one of the bearing state vibration signals in the test sample set is selected, and the denoising effect of the bearing state number obtained by the above different methods is tested, and the test results are shown in Fig. 3.

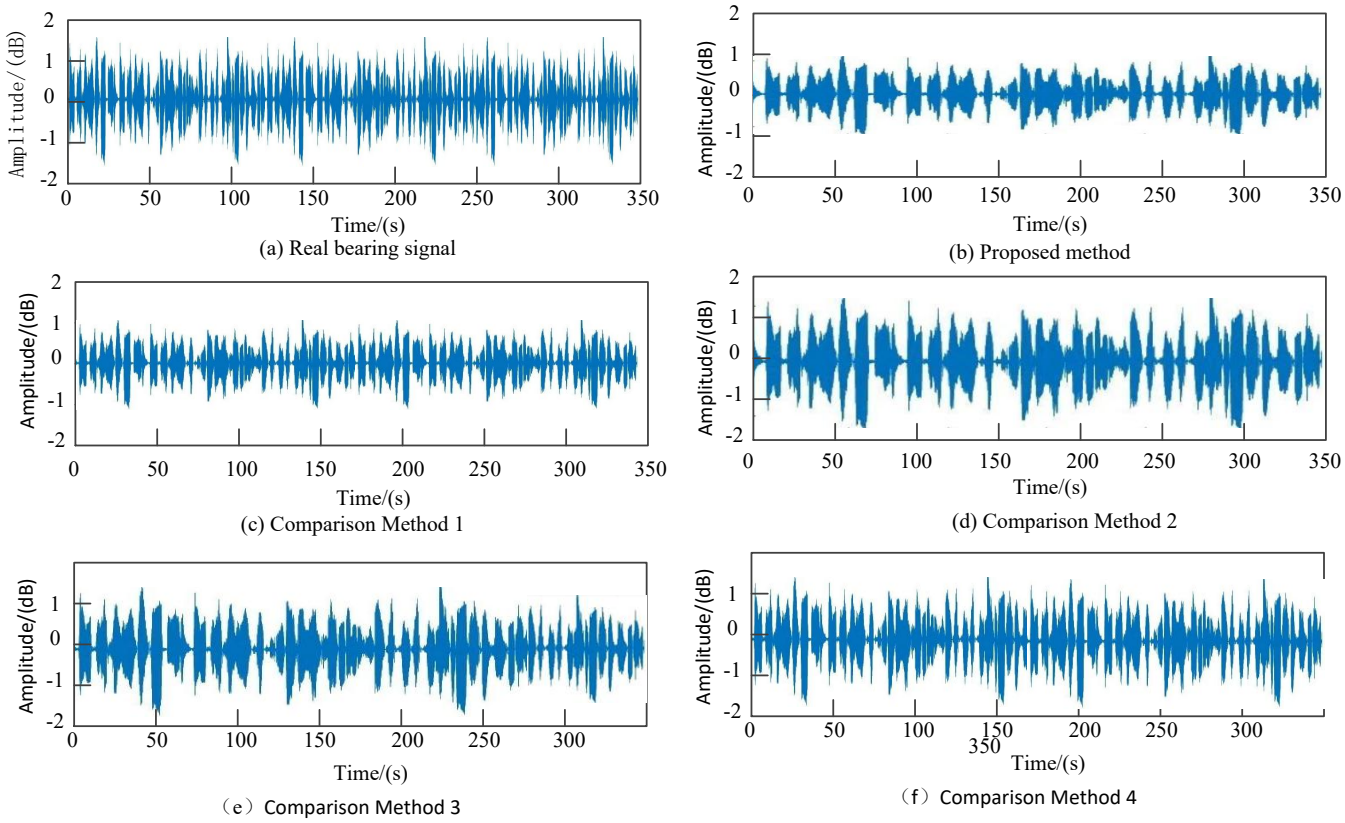

Figure 3. Test results of de-noising effect of bearing condition data by different methods.

Analysis of Fig. 3 shows that, to carry out bearing raceway wear monitoring, comparison method 1, although to a certain extent to reduce the noise of the vibration data signal, but the denoising process of selecting the threshold function there are problems, ignoring the preservation of the signal time-frequency information, so the method of the vibration data signal denoising, lost part of the normal signal in the frequency domain; Comparison method 2 and method 3, the denoising process ignores the continuity of the signal, so when the two methods in the denoising process, denoising effect is significantly lower than the proposed method denoising effect; Comparison method 4 due to the bearing fault diagnosis, did not design a targeted vibration data signal denoising content, so the method of denoising effect is not satisfactory; The proposed method is based on the traditional soft and hard threshold function structure, and improves a modified wavelet modal maxima denoising algorithm to denoise the collected vibration signals, so the method not only strips the noise in the signals completely, but also preserves the internal information of the normal signals. It can be proved that the proposed method is effective for bearing raceway wear monitoring.

4.2. Bearing raceway wear category identification accuracy test

When using the proposed method, comparing method 1, comparing method 2, comparing method 3, comparing method 4 to carry out bearing wear fault monitoring, and then test the sample set, select 200 test samples, construct 5 groups of sample data, each group of data corresponds to a wear fault type, the specific bearing wear types of the above five methods are identified, and the identification effect is shown in Table 3.

Table 3. Test results of accuracy in identifying bearing wear categories using different methods.

Test sample group	Actual type of bearing wear	Accuracy test results of bearing wear category recognition				
		Proposed method	Comparison method1	Comparison method2	Comparison method3	Comparison method4
1	Abrasive wear	Abrasive wear	Abrasive wear	Abrasive wear	Abrasive wear	Abrasive wear
2	Fatigue wear	Fatigue wear	Fatigue wear	Fatigue wear	Fatigue wear	Adhesive wear
3	Abrasive wear	Abrasive wear	Adhesive wear	Abrasive wear	Abrasive wear	Corrosive wear
4	Adhesive wear	Adhesive wear	Adhesive wear	Corrosive wear	Adhesive wear	Corrosive wear
5	Corrosive wear	Corrosive wear	Corrosive wear	Adhesive wear	Adhesive wear	Corrosive wear

Analysis of Table 3 shows that the proposed method can accurately identify the wear type of bearing raceway when carrying out bearing wear monitoring, mainly because the proposed method uses the clustering algorithm to interpolate the missing values of the collected bearing vibration data signals in advance to complete the bearing vibration data signals when monitoring the wear damage. Therefore, when the proposed method is utilized for bearing raceway wear monitoring, the accuracy of wear fault monitoring is better.

4.3. Actual bearing raceway wear practical effect test

Based on the above experimental results, we continue to use the above five methods to carry out raceway wear monitoring of bearings, and test the actual monitoring effect of different methods, and the test results are shown in Fig. 4.

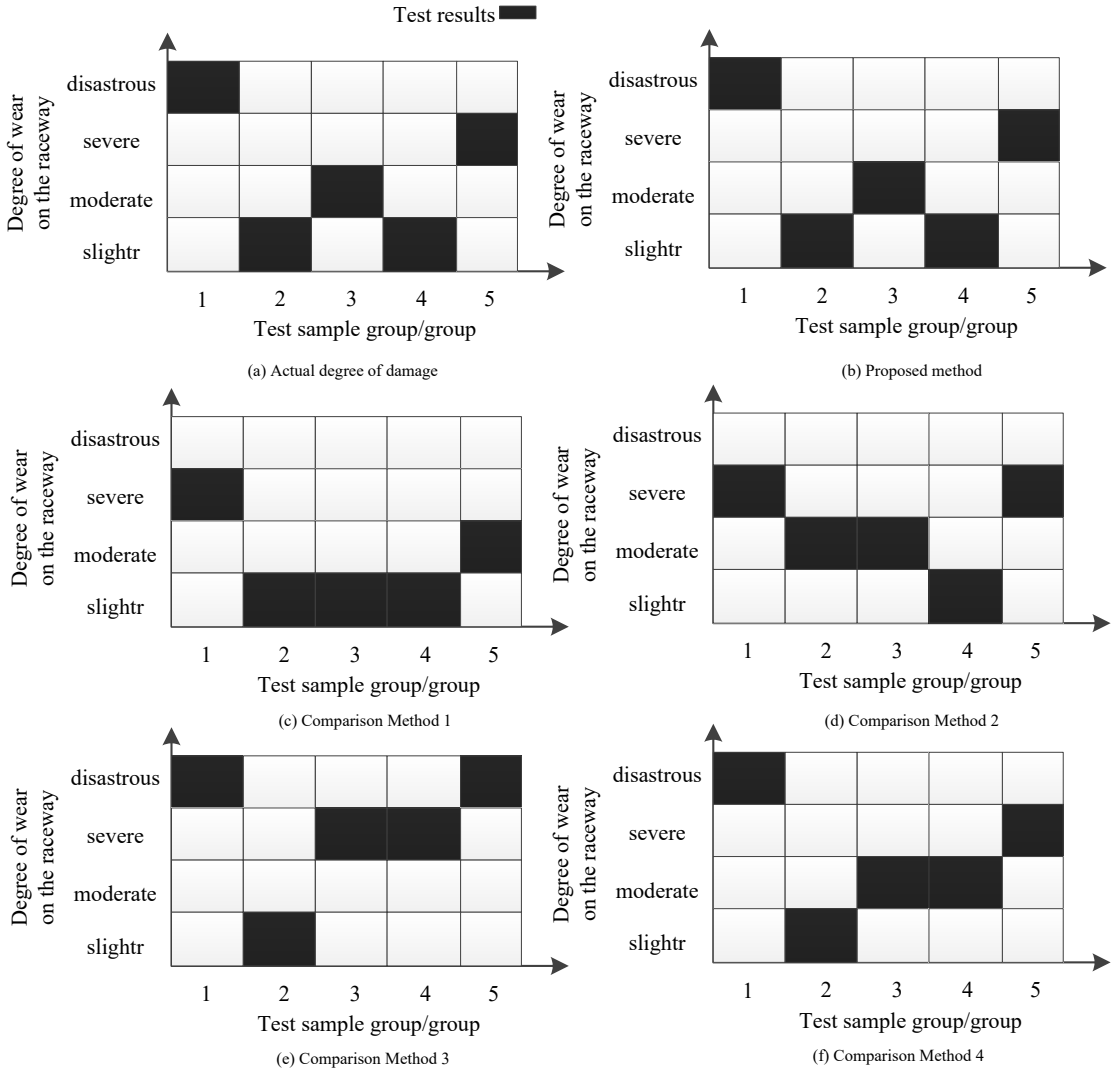

Figure 4. Test results of actual wear monitoring effectiveness using different methods.

Analyzing Fig. 4, it can be seen that when carrying out bearing rolling wear monitoring, comparing method 1 due to the serious redundancy of the algorithm parameters in the Fourier transform of the current signal, the performance of this method is lower in fault monitoring; comparing method 2 due to the extraction of bearing acoustic pattern characteristics, ignoring the interference of the external influencing factors of the bearing operation, so the method is lower in the monitoring accuracy when monitoring the bearing wear faults; comparing method 3 due to the introduction of the migration learning principle, migration learning parameters and its own network structure before the mismatch problem, so the method in the bearing wear monitoring, the monitoring effect is poor; comparison method 4 due to the collection of state data signal, the process of a large number of noise data signal, so the method in the bearing wear monitoring, the detection effect is not ideal; and the proposed method that before the raceway wear fault monitoring, aiming at the temperature data collected for the use of bearings, the health status of bearings is effectively identified. Therefore, this method has a good monitoring effect in the bearing wear monitoring.

5. Conclusion

As the main component of wind turbine power generation, it is important to monitor the raceway wear of the yaw bearing in wind turbines. Aiming at the problems existing in the traditional monitoring methods, an algorithm is proposed to monitor the large-area deep wear failure of the raceway of the yaw bearing in wind turbines. The method firstly collects the bearing state data and implements effective denoising process to increase the data quality; utilizes sub-domain adaptive adversarial migration network to establish a model for unhealthy bearings to carry out wear and failure feature extraction and classification process; finally, based on the wear warning mechanism, real-time monitoring of wear and tear to realize the large-area deep-wear monitoring of the bearing raceway.

References

Roga S, Bardhan S, Kumar Y, et al. Recent technology and challenges of wind energy generation: A review. Sustainable Energy Technologies and Assessments, 2022, 52: 102239. https://doi.org/10.1016/j.seta.2022.102239

Pryor S C, Barthelmie R J, Bukovsky M S, et al. Climate change impacts on wind power generation. Nature Reviews Earth & Environment, 2020, 1(12): 627-643. https://doi.org/10.1038/s43017-020-0101-7

Xu J, Benson S, Wetenhall B. Comparative analysis of fatigue life of a wind turbine yaw bearing with different support foundations. Ocean Engineering, 2021, 235: 109293. https://doi.org/10.1016/j.oceaneng.2021.109293

Jin X, Chen Y, Wang L, et al. Failure prediction, monitoring and diagnosis methods for slewing bearings of large-scale wind turbine: A review. Measurement, 2021, 172: 108855. https://doi.org/10.1016/j.measurement.2020.108855

Xu B, Xie Z, Chen S, et al. Diagnosis method of rolling bearing outer raceway fault in induction motors. Electric Machines and Control, 2022, 26(4): 1-8. https://doi.org/10.15938/j.emc.2022.04.001

Hu L, Gong Y, Zhang Y, et al. Pumped storage unit bearing fault diagnosis with a combination of sound and vibration dual-modal. Power System Protection and Control, 2024, 52(11): 1-10. https://doi.org/10.19783/j.cnki.pspc.231422.

Li L, Ma Z, Yu Z, et al. Quantitative diagnosis method for bearing faults based on MTF and improved residual network. Journal of Northeastern University (Natural Science Edition), 2024, 45(5): 697-706. https://doi.org/10.12068/j.issn.1005-3026.2024.05.012

Zhao Y, Wang X, Han S, et al. Fault diagnosis for abnormal wear of rolling element bearing fusing oil debris monitoring. Sensors, 2023, 23(7): 3402. https://doi.org/10.3390/s23073402

Cheng Y, Chen L, Meng X, et al. Testing and analysis of the ultimate load of yaw bearings in a certain type of wind turbine. Mechanical and Electrical Engineering, 2024, 42(2): 247-256. https://doi.org/10.3969/j.issn.1001-4551.2025.02.006

Wu Y. Cause of outer ring fracture of yaw double row ball bearing in wind-power. Physical Testing and Chemical Analysis(Part A:Physical Testing), 2023, 59(1): 54-56. https://doi.org/10.11973/lhjy-wl202301015

Riaboff L, Shalloo L, Smeaton A F, et al. Predicting livestock behaviour using accelerometers: A systematic review of processing techniques for ruminant behaviour prediction from raw accelerometer data. Computers and Electronics in Agriculture, 2022, 192: 106610. https://doi.org/10.1016/j.compag.2021.106610

Li R, Ma Y. Visualization of self-healing data flow for temperature sensor communication module. Computer Simulation, 2021, 38(4): 134-138. https://doi.org/10.3969/j.issn.1006-9348.2021.04.026

Wang Z, Juhasz Z. GPU implementation of the improved CEEMDAN algorithm for fast and efficient EEG Time–frequency analysis. Sensors, 2023, 23(20): 8654. https://doi.org/10.3390/s23208654

Shen Z, Ban W. Machine learning model combined with CEEMDAN algorithm for monthly precipitation prediction. Earth Science Informatics, 2023, 16(2): 1821-1833. https://doi.org/10.1007/s12145-023-01011-w

Wang H, Liu Z, Peng D, et al. Attention-guided joint learning CNN with noise robustness for bearing fault diagnosis and vibration signal denoising. ISA Transactions, 2022, 128: 470-484. https://doi.org/10.1016/j.isatra.2021.11.028

Tripathi P M, Kumar A, Komaragiri R, et al. A review on computational methods for denoising and detecting ECG signals to detect cardiovascular diseases. Archives of Computational Methods in Engineering, 2022, 29(3): 1875-1914. https://doi.org/10.1007/s11831-021-09642-2

Cui B, Weng Y, Zhang N. A feature extraction and machine learning framework for bearing fault diagnosis. Renewable Energy, 2022, 191: 987-997. https://doi.org/10.1016/j.renene.2022.04.061

Li Y, Tang B, Jiang X, et al. Bearing fault feature extraction method based on GA-VMD and center frequency. Mathematical Problems in Engineering, 2022, 2022(1): 2058258. https://doi.org/10.1155/2022/2058258

Han X, Wu Y, Wan R. A method for style transfer from artistic images based on depth extraction generative adversarial network. Applied Sciences, 2023, 13(2): 867. https://doi.org/10.3390/app13020867

Lowney B, Lokmer I, O'Brien G S, et al. Pre-migration diffraction separation using generative adversarial networks. Geophysical Prospecting, 2021, 69(5): 949-967. https://doi.org/10.1111/1365-2478.13086