中國機械工程學刊第四十五卷第六期第 595~604 頁(民國一百一十三年)
Journal of the Chinese Society of Mechanical Engineers, Vol.45, No.6, pp595~604 (2024)

Early Fault Diagnosis for Wind Turbine Gearbox Based on Multi-Source Feature Fusion

Li Wang*, He Jianjun**, Wang Jiawei***, Tang Zhiwei***, Jie Jun****, Luo Guangwu****and Liu yan****

Keywords： Wind turbine Gearbox; Multi-source information coupling; KPCA; VMD

ABSTRACT

As the primary moving part of a wind turbine, the gearbox has a high failure rate and is particularly detrimental to the device. The diagnosis of early gearbox problem signals is less effective using the typical vibration detection techniques now in use. Considering this, Based on the KPCA-VMD approach, this research offers a wind turbine gearbox early fault monitoring and multidimensional feature assessment method for analyzing wind turbine gearbox inconspicuous early failure signals. Firstly, the pre-processed dataset is subjected to feature extraction, the gearbox feature data is downscaled and reconstructed by the KPCA method, the gearbox status is monitored using two statistics, T^2 and SPE, and the monitored abnormal signals are analysed by VMD. The experimental data show that the method can effectively diagnose the gear early failure characteristic frequency.

INTRODUCTION

Gearboxes, as the core connecting and driving mechanical equipment of wind turbines, are characterized by precise and complex mechanical structures (Salameh et al., 2018). Due to the long-term work at high altitude on 100 meters, it is constantly subjected to changing wind speed loads and mechanical loads (Guo and Sheng et al., 2020) and when cracks or even damage behaviors occur in the variable speed gearbox, the crew often cannot detect them earlier, which leads to a decrease in equipment reliability and safety. (Yang et al., 2023; Shanbr et al., 2018) Therefore, for how to gearbox early fault monitoring and diagnosis, health evaluation, effective extraction of early fault characteristics and accurate condition monitoring is an important direction of research.

Wind turbine component damage monitoring, generally based on monitoring data in the SCADA system, such as main bearing temperature, gearbox temperature, oil temperature, etc., has the defects of constant data, small availability and many missing values. (Zhu et al., 2019; Igba et al., 2015) The other is a wind turbine condition monitoring system system (CMS) that monitors sensor signals such as vibration or displacement, but is still primarily based on a single signal source for early warning analysis. (Entezami et al., 2012; Dao et al., 2022) This routine monitoring, operation, and maintenance fail to make full use of existing technical methods to accurately assess early failures of wind turbine gearboxes, so there is huge room for technical improvement.

The signal time domain characteristics are generally periodic and smooth when the gear components are in normal operation, and the signal frequency characteristics will include the rotation frequency of each shaft and the meshing frequency of large and small gears. (Wang et al., 2023; Zhang et al., 2022; Pan et al., 2019) When the damage failure of the component occurs, the fault signal will appear shock or modulation phenomenon, the fault signal characteristics are particularly significant when the broken tooth or tooth surface peeling. (Touti et al., 2023; Tong et al., 2019) However, the early signal often has non-linear, non-smooth, low amplitude, low signal-to-noise ratio, and other characteristics, the fault characteristics are hidden, and feature extraction is difficult. (Yaghoubi et al., 2022) Vikas Sharma et al. (Sharma et al., 2016) reviewed various time-domain state indicators for wind turbine gear fault diagnosis. Jong et al. (Jong et al., 2012) analyzed wind turbine

Paper Received April, 2024. Revised May, 2024. Accepted June, 2024. Author for Correspondence: Jianjun He.

* *Engineer, CHN Energy Hunan Power Co., Ltd, ChangSha 410000, China.*

** *Professor, School of Energy and Power Engineering, Changsha University of Science and Technology, Changsha 410114, China;*

*** *Graduate Student, School of Energy and Power Engineering, Changsha University of Science and Technology, Changsha 410114, China;*

**** *Engineer, Longyuan Jiangyong Wind Power Generation Co., Ltd of CHN Energy Group, Changsha 410000, China.*

planetary gearbox condition monitoring using time-synchronous mean autocorrelation. Abboud et al. (Abboud et al., 2017) implemented fault diagnosis for wind turbines under various operating conditions using envelope spectra, etc. However, the above methods suffer from problems such as modal confounding and endpoint effects when applied to the actual complex signal analysis.

Variational Mode Decomposition (VMD) is an adaptive signal processing method based on Wiener filtering, which decomposes the original signal into a finite number of IMFs in different frequency bands, overcoming endpoint effects and modal aliasing problems. (Chen et al., 2019; Yan et al., 2019; Ren et al., 2019) The core of the VMD algorithm is the construction and solution of the variational problem.

Kernel Principal Component Analysis (KPCA) is a nonlinear feature extraction method, (Shen et al., 2022) which can well eliminate redundant correlations between data and extract nonlinear features that retain the main information. Based on this multivariate statistical method of KPCA, the frequent detection method is able to improve the computing rate of the computer, and at the same time, it is also able to identify and judge the early faults of the gearbox system. (Liu et al., 2023; Navi et al., 2015; Pacheco-Chérrez.et.al., 2022)

However, the use of "cliff" metrics alone to filter critical IMFs can easily result in missing early feature information or misdiagnosis of faults. (Li et al., 2020) The study of VMD for noise reduction of signals and screening of key IMFs will be of great importance for the fine diagnosis of gearbox faults and for early fault diagnosis. (Zheng et al., 2023) Currently, there is no mature method for monitoring the early fault signals of gearboxes, this paper proposes a wind power gearbox early fault monitoring and multidimensional feature evaluation method based on the KPCA-VMD method, which can effectively identify and diagnose the early fault signals of gearboxes, so as to realise the on-line monitoring of the early faults of the gearboxes, to prevent further deterioration of the gearboxes, and to safeguard the safety of the operation of wind turbines.

METHODOLOGY

Wind turbine operating conditions are complex and fault-prone, and monitoring methods based on a single signal are often insufficient for accurate fault diagnosis. By monitoring multi-source information points, fusing data features in multiple dimensions, and then downscaling for diagnosis, the accuracy of mechanical fault diagnosis will be further improved, providing a stronger guarantee for the safe operation of wind turbines.

As shown in Fig.1, this paper will collect multi-source sensor information, using the gearbox of a wind turbine in a wind farm as a signal source. Vibration, acoustic emission, displacement, and speed sensors are installed to monitor the real-time changing vibration and speed signals to effectively monitor and analyze the gear meshing frequency and impact energy. The collector aggregates these data and transmits them from the fiber optic switch to the server's database for later processing and research of the data.

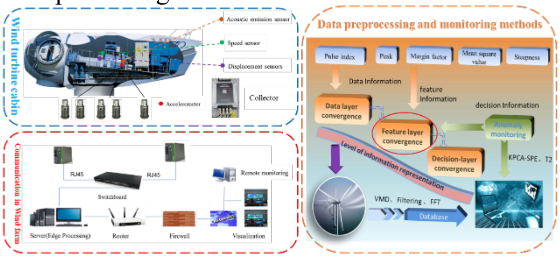

Fig. 1. Framework diagram of wind turbine gearbox early fault abnormal monitoring and diagnosis system

Besides, the KPCA method is used to achieve multi-dimensional data fusion at the feature layer to achieve the purpose of abnormal state monitoring of wind turbine gearboxes. Compared with the traditional PCA method, KPCA can better handle complex nonlinear data and improve the accuracy and effectiveness of feature extraction. VMD is an adaptive signal processing method that can decompose different modal components in a signal, which is suitable for the analysis of non-stationary signals. Compared with the traditional Empirical Mode Decomposition (EMD), VMD realizes the adaptive optimization of modal decomposition through the variational model, which reduces the modal aliasing phenomenon and improves the accuracy of decomposition. The KPCA-VMD method is then used to de-noising, reconstruct, and diagnose faults in the abnormal data, laying the foundation for intelligent safe operation and health evaluation of wind turbines.

Feature fusion of multi-source data

Data pre-processing

The fusion of multi-source data firstly requires pre-processing the original information, cleaning some missing values, and standard and continuous data to provide for data mining and training later. The quality of the data will directly affect the results of signal monitoring. We will perform singularity removal based on the 3σ criterion. The standard deviation σ is first obtained according to the Bessel formula, a threshold of 3σ is set, and then the mean value of a certain set of data is calculated. The value of each data's deviation from the mean is compared by executing a judgment statement, and if the result is greater than 3σ, the outlier is removed and vice versa.

Secondly, the signal data needs to be de-noised. The excitation force is applied to the gearbox and the gears produce vibration, which is transmitted to the gear train, shaft, bearing housing, and box, mechanical vibration in the air produces noise, so the signal data often contains the characteristic frequency, the intrinsic frequency, noise coexistence of complex signal information. The noise signal frequency is high and the signal feature peak frequency is generally dominated by the low-frequency component. Therefore, before realizing the signal feature extraction, the high-frequency part should be processed with noise reduction using methods such as KPCA, VMD, etc, so as to eliminate or reduce the modulation effect of the high-frequency signal on the low-frequency signal caused by noise, or the failure to produce the modulation phenomenon of the low-frequency signal on the high-frequency signal.

Feature extraction and fusion

Feature extraction and fusion means transforming data signals from different types of sensors of the same measurement point into a uniformly structured data expression, rowing them into feature vectors for identification, and then analyzing the normalized data to complete the analysis and evaluation of the object being measured. As shown in the Table.1 below, the maximum, minimum, and peak values of the displacement signal; the average signal level (ASL) and root mean square (RMS) of the AE signal; and the stiffness, peak indicator, pulse indicator, and margin indicator of the vibration signal are extracted.

Table 1. Expressions for characteristic parameters

Sensor Signal type	Characteristic indicators	function expression
Displacement signal	maximum	$x_{max}=max\{x_i\}$
	minimum	$x_{min}=min\{x_i\}$
	peak	$x_{p\text{-}p}=X_{max}\text{-}X_{max}$
Acoustic emission signal	Average signal level	$x'=\frac{1}{N}\sum_{i=0}^{N-1}\lvert x_i\rvert$
	Root mean square	$x_r=(\frac{1}{N}\sum_{i=0}^{n-1}\lvert x_i\rvert^{\frac{1}{2}})^2$
Vibration signal	stiffness	$x_q=(\frac{1}{N}\sum_{i=0}^{N-1}{x_i}^4/{x_a}^2)-3$
	peak index	$C=\frac{x_p}{x_{rms}}$
	pulse index	$I=\frac{x_p}{x'}$
	margin index	$L=\frac{x_p}{x_r}$

Z-score standardization: converts the original data into a standard normal distribution, x_j a distribution with a mean of $\bar{x}$ and a standard deviation of δ. The expression is shown below:

$$x_j{}^*=\frac{x_j-\bar{x}}{\delta} \tag{1}$$

Data feature dimensionality reduction and anomaly monitoring based on the KPCA method

The KPCA model for data feature dimensionality reduction

Assume that the training sample for the gearbox multi-source data model is shown below:

$$\mathrm{T}=\begin{bmatrix} \mathrm{x}_1(1) & \mathrm{x}_1(2) & \mathrm{x}_1(3) & \dots & \mathrm{x}_1(m) \\ \mathrm{x}_2(1) & \mathrm{x}_2(2) & \mathrm{x}_2(3) & \dots & \mathrm{x}_2(m) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathrm{x}_\mathrm{n}(1) & \mathrm{x}_\mathrm{n}(2) & \mathrm{x}_\mathrm{n}(3) & \cdots & \mathrm{x}_\mathrm{n}(m) \end{bmatrix}_{m\times n} \tag{2}$$

The m evaluation samples, with n parameter indicators, are mapped to a high-dimensional feature space via a non-linear mapping($\phi: \mathrm{R}^n \rightarrow \mathrm{F}$) , followed by the mapped dataset ($\phi(\mathrm{x}) = \{\phi(x_1), \phi(x_2), \dots, \phi(x_n)\}$) being centered so that it satisfies the following equation:

$$\frac{1}{N}\sum_{i=1}^{n}\phi(x_i) = 0 \tag{3}$$

The covariance matrix of the sample C^F is as follows:

$$C^F = \frac{1}{N}\sum_{\mathrm{i}=1}^{\mathrm{n}}\phi(\mathrm{x_i})\phi(\mathrm{x_i})^T \tag{4}$$

The covariance characteristics are decomposed as:

$$\lambda\omega=C^F\omega \tag{5}$$

where λ is the eigenvalue of the covariance matrix, ω is its eigenvector, $\omega = \sum_{\mathrm{j}=1}^{\mathrm{n}}\alpha_i\phi(\mathrm{x_j})$, and α_i is the correlation coefficient.

To solve the nonlinear map $\phi(\mathrm{x})$, we introduce the kernel function K so that each element in the $\phi(\mathrm{x})$ satisfies the following equation:

$$[K]_{ij} =(\phi(x_i), \phi(x_j)) \tag{6}$$

After the data is centralized and simplified in high-dimensional feature space, the sample data principal t_k can be calculated:

$$t_k =(\omega_k, \phi(\mathrm{x}))= \sum_{\mathrm{i}=1}^{\mathrm{n}}\alpha_i^k \cdot k(x_i, x) \tag{7}$$

The kernel function uses the commonly used Gaussian radial kernel function, as shown below:

$$k(x, y) = e^{-\frac{\|x-y\|}{c}} \tag{8}$$

where c is the kernel parameter. The method of p_0 the number of principal elements is obtained from the cumulative variance.

The KPCA-SPE/T^2 value for data anomaly monitoring

The Squared prediction error (SPE) statistic is a measure of the degree of deviation from the kernel principal model, reflecting the process by which the data deviate from the normal correlation. After nonlinearly mapping to high-dimensional space can be expressed in the following expression:

$$\mathrm{SPE}=\sum_{k=1}^{p}({t_k^j}^2)\text{-}\ \sum_{k=1}^{\gamma}({t_k^j}^2) \tag{9}$$

The control limit solution formula is as follows:

$$\mathrm{SPE}_{lim} = \mathrm{g}\chi_{h,\alpha}^2 \tag{10}$$

where $\mathrm{g}=\frac{b}{2a}$ is the weighted parameter, $\chi_{h,\alpha}^2$ is the χ^2 distribution with the confidence of α and degrees of freedom h, and a,b is the mean and variance under normal working conditions.

The Hotelling statistic (T^2) characterizes the trend in the sample and test data in the primary metric space and the degree of deviation in the model and is a reflection of the variation within the KPCA model. Its equation is shown below:

$$T^2 = [t_1, t_2, \dots, t_p]\Lambda^{-1}[t_1, t_2, \dots, t_n]^T \quad (11)$$

Where t_p is the principal element of the high-dimensional feature space; p is the number of principal elements; and Λ is the diagonal matrix formed by the eigenvalues corresponding to the components of the kernel principal element.

The control limit for the T^2 threshold can be obtained from the F distribution：

$$T_{lim}^2 = \frac{l(n^2-1)}{n(n-1)} F_\alpha(l, n-l) \quad (12)$$

Where $F_\alpha(l, n-l)$ corresponds to the critical value of the F distribution at a test level of α and degrees of freedom of l, n-l.

Diagnosis of gearbox damage behavior based on the KPCA-VMD method

The principles of the VMD method

The objective of the VMD method is to decompose the signal into eigenmode components (IMF) of different frequency bands, each IMF mode being closely distributed around the central frequency ω_k. The number of IMFs is optimally chosen to set the number of IMFs K. A suitable value of K can avoid frequency aliasing during the decomposition of the variational modes.

The variational problem is constructed as follows:

$$\begin{cases} \min\limits_{\{u_k\},\{\omega_k\}} \left\{\sum_k \left\| \partial_t \left[\left(\delta_t(t) + \frac{j}{\pi t}\right) * u_k(t)\right] e^{-j\omega_k t} \right\|_2^2 \right\}; \\ s.t. \sum_k u_k = f(t). \end{cases} \quad (13)$$

where k denotes the number of intrinsic mode components; $\{u_k\}$, $\{\omega_k\}$ denote the kth component and its center frequency, respectively; ∂_t denotes the gradient of the demodulated signal; $\delta_t(t)$ denotes the Dirac function; $*$ denotes the convolution operator; f(t) denotes the original signal.

The penalty factor α_1 and the Lagrange multiplier λ_1 are introduced to turn it into an unconstrained variational problem. The augmented Lagrange expression is given by:

$$L(\{u_k\}, \{\omega_k\}, \lambda_c) = \alpha_1 \sum_k \left\| \partial_t \left[\left(\delta_t(t) + \frac{j}{\pi t}\right) * u_k(t)\right] e^{-j\omega_k t} \right\|_2^2 + \left\| x_c(t) - \sum_k u_{k,c}(t) \right\|_2^2 + \langle \lambda_1(t), f(t) - \sum_k u_k(t) \rangle \quad (14)$$

The optimal solution is obtained by iteratively updating the values of u_k, ω_k, and λ by the alternating direction multiplier method to solve for the minimum value. The optimal solution is the intrinsic mode component $\{u_i\}$ and the respective central frequency $\{\omega_i\}$.

Noise reduction, Characterization, and comprehensive evaluation guidelines

A joint noise reduction strategy is proposed to extract weak fault features from the early fault signals of wind turbine gearboxes, which are non-linear, non-smooth, low amplitude, and low signal-to-noise ratio. Firstly, the KPCA method is used to de-noise and reconstruct the signals, and the signal-to-noise ratio (SNR) and root mean square error (RMSE) is used to quantify the noise reduction effects of different methods. Secondly, the IMFs of different frequency bands are obtained by VMD decomposition of the noise reduction data. Then, a "kernel principal" evaluation model based on multidimensional feature fusion is constructed. The larger the kernel principal component, the better the IMF can characterize the original signal, thus screening out the key IMFs that characterize the early faults. The reconstructed secondary noise reduction signal can further highlight the fault characteristics and is suitable for early fault diagnosis, as shown in Fig.2.

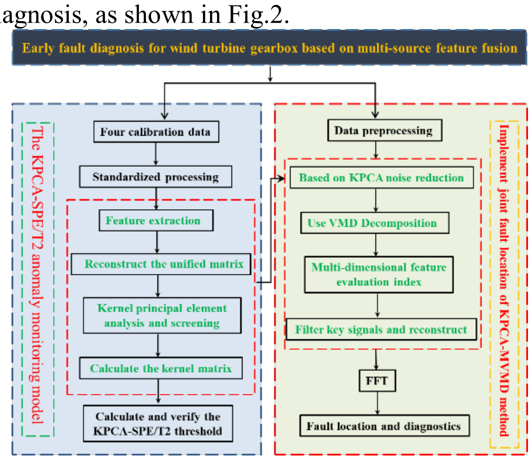

Fig. 2. Flow chart of early fault monitoring and diagnosis methods for wind turbines

CASE APPLICATION AND ANALYSIS

Data from the gearbox of a grid-connected wind farm were used to verify the validity of the proposed method. The gearbox has a primary parallel wheel with 29 teeth and a large gear with 100 teeth, a secondary pinion with 36 teeth and a large gear with 90 teeth, a sun wheel with 28 teeth, four planetary wheels with 36 teeth, and a ring with 100 teeth. Table 2 lists the details. The signal sampling frequency is 8000 Hz. The data collected was initially labeled by the wind farm technicians. Based on years of fault records from the wind farm, the operating conditions of the wind turbine gearboxes were classified into 4 categories according to the severity of the fault, i.e. normal, cracked, uniformly worn, and broken gears, with 65 samples for each of the 4 operating conditions and a single sample length of 1024.

Table 2. Gearbox gear information parameters

Gear type	Teeth	Quantity
primary parallel wheel	29	1
large gear	100	1
secondary pinion	36	1
large gear	90	1
sun wheel	28	1
planetary wheel	36	4
ring	100	1

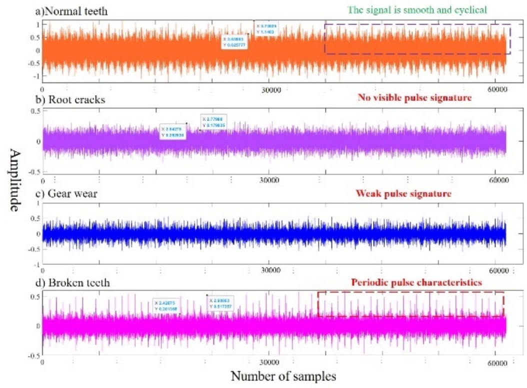

Fig. 3. Vibration signals in four operating states

As shown in Fig.3, the time domain signal amplitude of normal gears and tooth root cracks is relatively smooth and periodic in nature. Due to the weakness of the fault signal as well as noise interference, it is difficult to detect a periodic shock characteristic from the time domain amplitude even if an early gear fault occurs. In contrast to a serious fault phenomenon such as a broken tooth, it can be found in the time domain as a regular shock-type vibration with a sudden increase in amplitude, the frequency of the shock being equal to the rotational frequency of the shaft where the broken tooth is located.

Feature analysis

Due to the complexity of the signals, a single sensor feature cannot fully monitor the health status of the gearbox. The multi-source data sets of gears in four different states are pre-processed to extract the key features of different source signals separately to constitute multi-dimensional evaluation indexes, as shown in Table 1, to obtain a feature matrix of 9*132 for each group. Three features that are more sensitive to the gearbox fault state, namely, the margin indicator of the vibration signal, the peak indicator of the displacement signal, and the RMS indicator of the AE signal, are selected for visualization and analysis.

As shown in Fig.4a**)**, the three indicators are more obvious for the distinction between the normal state and serious gear failure state. However, the three states of normal teeth, cracked teeth, and uniformly worn teeth have the phenomenon of indicator amplitude overlap, which is not ideal for the early fault identification of gearboxes. This is because the occurrence of damage failure of gearbox system components generally starts from crack sprouting and then extends to fatigue, pitting, oxidation intensification, wear-off, etc. This process of crack expansion often goes through a series of stages.

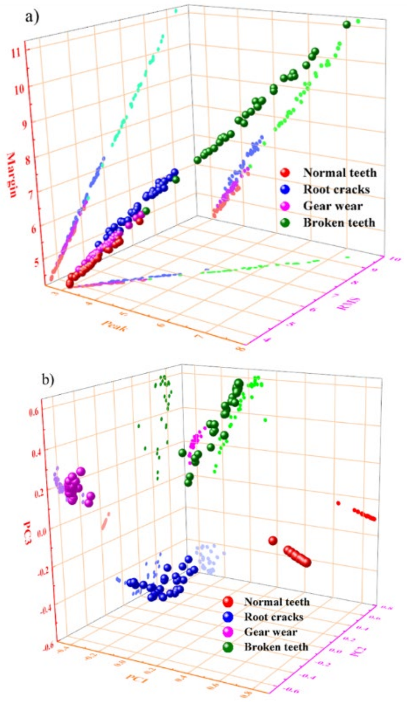

Fig. 4. Multi-source data feature visualization.
a) Before feature fusion dimensionality reduction;
b) After feature fusion dimensionality reduction.

The 9 multidimensional features are normalized to construct a unified judgment matrix, then subjected to KPCA dimensionality reduction. The Gaussian function is selected as the kernel function, and the different parameter gamma values of the kernel function will make the spatial distribution of the data separate. When the gamma value is too small, the model will be under-fitting; when the gamma value is too large, the model will be over-fitting. Therefore, the gamma values were tested in steps of 0.01 from the interval (0, 15), and the explained variance was used as a measure. The optimal gamma value of the model was found to be 0.13, and the explained variance was 0.55629. As shown in Fig.4b**)**, after the KPCA dimensionality reduction process, the data sets between the gear states achieve obvious clustering and state separability. Whether it is mapped in the PC1-PC2 plane or PC2-PC3 plane, the early root fault state characteristics are obviously distributed in different spaces from other state features, and there is no aliasing of the monitoring values. Thus, this model can successfully monitor gear outliers.

Abnormal monitoring of damage behavior of

gearbox components

The PC2 principal element value with better state classification is substituted into Eq. (11) to obtain the T^2 statistic as shown below. Fig.5b**)** shows the threshold limit of the T^2 statistic obtained by training the master metadata set of normal gears, and the limit is 7.97. Because the confidence limit of the threshold calculation is set to 0.95, it is reasonable for some test points to exceed the threshold limit, which is not contrary to the state of normal gears.

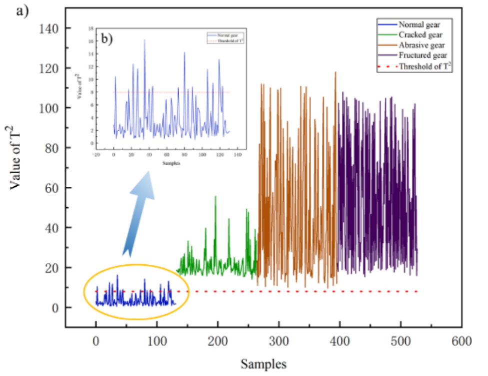

Fig. 5. Abnormal condition monitoring based on KPCA-T^2 values.
a) Gearbox T^2 threshold limit and each state test set;
b) Gearbox T^2 training set with threshold limit.

As shown in Fig. 5a**)**, when the gear state fails, the T^2 statistics of each of its data sets, exceed the red threshold line. At the same time, the T^2 magnitudes of the early fault data belong to different spaces compared with the severe fault occurrence data, and their distribution patterns are also consistent with the mapping pattern in Fig. 4b**)**. Therefore, it indicates that the reduced T^2 statistics can distinguish the early fault data exactly and effectively.

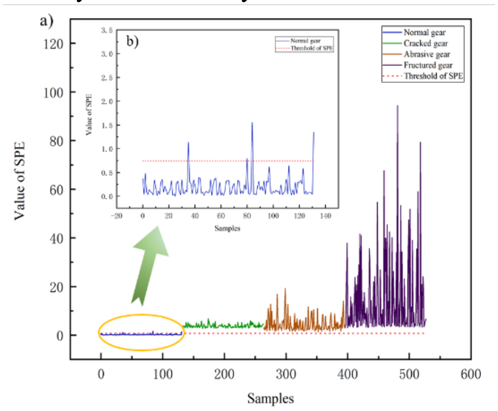

Fig. 6. Abnormal condition monitoring based on KPCA-SPE values.
a) Gearbox SPE threshold limit and each state test set;
b) Gearbox SPE training set with threshold limit.

The results of the SPE statistic as shown in Fig. 6. Compared with the T^2 statistic, the SPE statistic has fewer points above the threshold limit, and the value of the SPE statistic for normal gears tends to be smoother. For the three types of faulty gears, namely, cracked root gears, uniformly worn gears, and broken gears, the values of SPE statistics are also above the threshold limits, and both T^2 and SPE statistics are able to alarm the three types of faulty gears, and there are no missed and false alarms. By combining the T^2 and SPE statistics to monitor the gears, the status of the gears can be monitored more accurately.

Fault diagnosis and analysis of gearbox components

When the abnormal signal of the gearbox is monitored, the following is to further do fault diagnosis analysis for the gearbox components. Firstly, we need to determine the penalty function in the VMD method and set the appropriate number of IMFs K. The appropriate K value decomposition can effectively avoid the endpoint effect and modal confounding. The VMD parameters are determined by using the method proposed in the literature, and the average value of each state is sought 10 times to obtain the modal number K of normal gears as 6 and the penalty factor α as 2983; the modal number K of tooth root cracks as 6 and the penalty factor α as 3465. The judgment accuracy ε and noise tolerance τ have little influence on the decomposition results and are taken as default values. Fig. 7 shows the time domain diagram of each IMF of the gear signal after VMD, and Fig. 8 shows the frequency domain diagram of each IMF. From the figure, it can be seen that the frequency range of each IMF does not overlap, and there is no frequency mixing phenomenon.

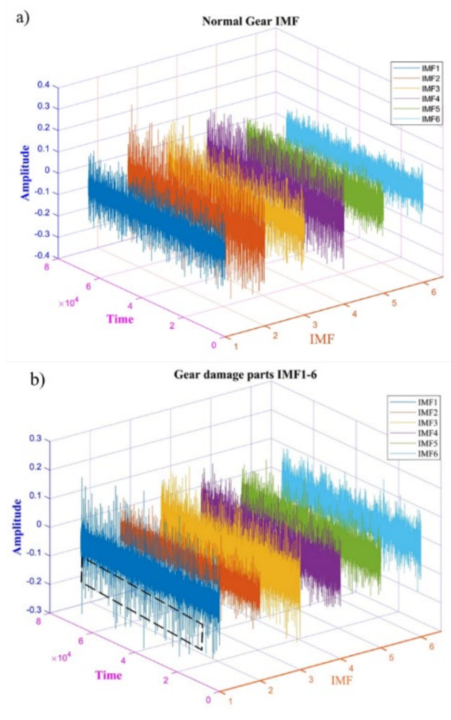

Fig. 7. IMF component time domain plot.
a) Normal gear;
b) Root cracked teeth.

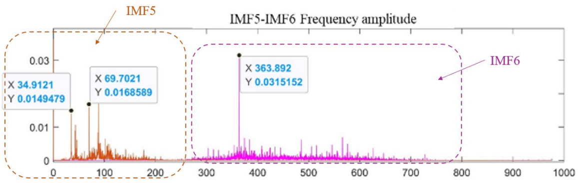

Fig. 8. IMF component frequency domain diagram.

In comparison with Fig. 7a), and Fig. 7b) although the increase in amplitude due to the increase in impact energy can also be found in the time domain, it is relatively not very obvious and the IMF variation pattern is not uniform. As shown in Fig.8, although the frequency diagram of the early fault also reads the engagement frequency characteristic (363.8 HZ), the marginal spectrum frequency is not obvious enough to diagnose that a fault has occurred at this moment. Therefore, it cannot be used as a definite conclusion to diagnose the cause of the early failure.

There are two key characteristics of wind turbine gearbox fault signals: impulsivity and cyclic smoothness. Existing studies usually consider only the impulsivity of the fault, and assessing the impulsivity of the fault by the stiffness alone can ignore some characteristic information of the IMF. The margin factor is more sensitive to the weak change of signal because it can well identify the early damage behavior of the component, and the envelope spectrum stiffness can comprehensively evaluate the cyclic smoothness of each IMF. As shown in Table 3, the multidimensional feature evaluation matrix of the "steepness-envelope spectral steepness-margin factor" of each IMF is calculated to screen for obvious early fault features.

Table 3. IMF components margin factor, steepness and envelope spectral steepness

Modal component	Margin factor	Steepness	Envelope spectral steepness
IMF1	0.3902	2.635	3.246
IMF2	0.4123	3.075	2.603
IMF3	**0.3900**	**4.331**	3.973
IMF4	0.5163	3.264	4.132
IMF5	0.6826	3.753	5.949
IMF6	**0.6833**	**5.832**	**6.977**

From Table 3, we can see that IMF5 and IMF6 contain rich fault information, and the reconstructed signals are shown in Fig. 9b) by using IMF5 and IMF6 as the key components. If the traditional unidimensional index "stiffness" is used as the selection principle of key components, IMF3 and IMF6 are reconstructed, and the reconstructed signal is shown in Fig.9a). It is found that the reconstructed signal in Fig.9b) shows a higher amplitude, and more periodic shock pulses, and the fault characteristics are more obvious.

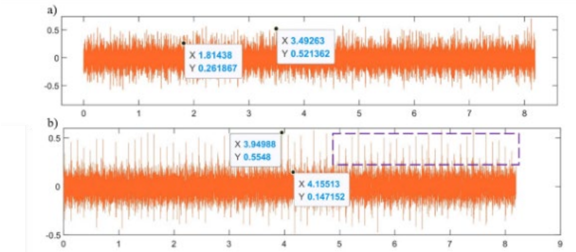

Fig. 9. Comparison chart of the filtered IMF reconstruction signal.
a) Traditional screening;
b) Multidimensional feature evaluation screening.

The gear meshing frequency component (363.8Hz) of the reconstructed signal in Fig.9b) is extracted with the fault frequency characteristics. The characteristic frequency formula can be calculated to obtain the planetary shaft rotation frequency of 10.15Hz and the medium speed shaft rotation frequency of about 4.06Hz, etc. Compared with Fig.8, it is clearly observed from Fig.10 that not only the low-speed shaft (4Hz) rotational frequency modulation side band is obviously present on both sides of the faulty gear meshing frequency, but also the medium-speed shaft rotational frequency modulation (10Hz) sideband is well revealed, i.e. 351.6 and 373.4Hz. The results show that the KPCA-VMD-based method can effectively diagnose and analyze the early fault characteristic frequency of gears, and diagnose the signal modulation phenomenon due to the early fault of gears in both low-speed and medium-speed shafts.

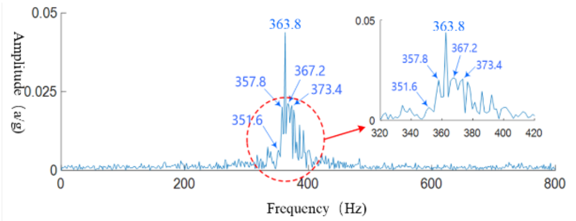

Fig. 10. Spectrogram of the reconstructed signal.

In order to further verify the effectiveness of the KPCA-VMD method for early gear fault analysis, the noise reduction effects of different methods were compared, as shown in Fig.11. The signal was first de-noised and reconstructed using the KPCA method, followed by noise reduction and reconstruction using the VMD method, followed by joint noise reduction using the KPCA-VMD method, and the noise reduction effect was analyzed by comparing the wavelet threshold noise reduction processing methods.

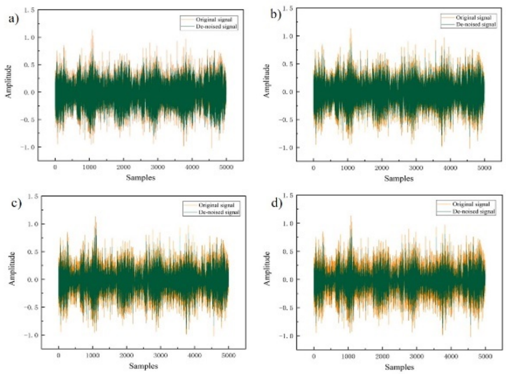

Fig. 11. Comparison chart of noise reduction effect of different methods.
a) KPCA;
b) VMD;
c) The Wavelet threshold noise reduction;
d) The KPCA-VMD method.

From the Fig.11, we can see that the green part is obviously reduced compared with the orange part, which indicates that the KPCA and VMD methods have some effect on the signal noise reduction, and the rising edge and falling edge of the reconstructed signal still retains a good part of the mutation, but the overall noise still exists, and the signal is not smooth enough. The wavelet threshold noise reduction was used, and the Daubechies wavelet basis function and 8-layer decomposition were selected. After several trials of setting the adjustment factor a=0.04, the noise reduction effect in the signal was improved, but compared with the joint KPCA-VMD noise reduction, the effect could not be directly judged from the graph.

Therefore, SNR and RMSE were used as evaluation indexes to quantitatively analyze the noise reduction effect of different methods, and the results are shown in Table 4. The SNR of the reconstructed signal is improved by 56.42% after the proposed method, and the SNR of the reconstructed signal is higher and the RMSE is lower than that of the wavelet threshold noise reduction.

Table 4. Comparison of indicators for different noise reduction strategies

Noise reduction methods	SNR	RMSE
KPCA	11.533	0.256
VMD	12.437	0.168
Wavelet de-noising	13.069	0.235
KPCA-VMD	**15.374**	**0.168**

CONCLUSION

To scientifically evaluate the gearbox health status, diagnose early faults, and improve the problem of inaccurate monitoring accuracy of a single signal source, this paper proposes a method for early fault monitoring and multidimensional feature evaluation of wind turbine gearboxes based on the KPCA-VMD method, which provides a new idea for the whole life cycle health assessment of wind turbine gearboxes. The main contributions are:

1) When multi-dimensional data sources are added to evaluate the early fault characteristics of different spatial segments, the KPCA method can be used to effectively reduce the data dimensionality, extract the key information of the data, and discover the hidden relationships between the data, to achieve the purpose of noise reduction, state classification visualization, and early fault monitoring.

2) Combined with engineering practice, the joint KPCA and VMD methods analyze the early inconspicuous fault impact signals, evaluate the early fault signals jointly with the margin, impulsivity, and cyclic smoothness indexes, and strengthen the early fault characterization ability of the feature components.

3) The KPCA-VMD method combines the advantages of KPCA's nonlinear feature extraction and VMD's modal decomposition, and shows stronger noise immunity and robustness in the face of noise data. KPCA effectively reduces dimensionality and removes noise, while VMD accurately separates the modal components of the signal.

The research results show that the method is applicable and superior to early fault monitoring and diagnosis, and can be used for predictive maintenance of wind turbine gearboxes.

ACKNOWLEDGMENT

The authors gratefully acknowledge financial support from the National Energy Group Technology Innovation Project (HJLFD).

REFERENCES

Abboud D, Antoni J, Sieg-Zieba S, et al., “Envelope analysis of rotating machine vibrations in variable speed conditions: A comprehensive treatment,”. *Mechanical Systems and Signal Processing*, Vol.84, pp.200-226 (2017).

Chen X, Yang Y, Cui Z, et al., “Vibration fault diagnosis of wind turbines based on variational mode decomposition and energy entropy,”. *Energy*, Vol.174, pp.1100-1109 (2019).

Dao P B., “Condition monitoring and fault diagnosis of wind turbines based on structural break detection in SCADA data,”. *Renewable Energy*, Vol.185, pp.641-654 (2022).

Entezami M, Hillmansen S, Weston P, et al., “Fault detection and diagnosis within a wind turbine mechanical braking system using condition monitoring,”. *Renewable Energy*, Vol.47, pp. 175-182 (2012).

Guo Y, Sheng S, Phillips C, et al., “A methodology for reliability assessment and prognosis of bearing axial cracking in wind turbine gearboxes,”. *Renewable and Sustainable Energy Reviews*, Vol.127, No.109888 (2020).

Ha, J. M., Youn, B. D, Oh, H, Han, B , Jung, Y. , & Park, J., "Autocorrelation-based time synchronous averaging for condition monitoring of planetary gearboxes in wind turbines,". *Mechanical Systems & Signal Processing*, pp.70–71(MAR.), pp.161-175 (2012).

Igba J, Alemzadeh K, Durugbo C, et al., "Performance assessment of wind turbine gearboxes using in-service data: Current approaches and future trends,". *Renewable and Sustainable Energy Reviews*, Vol.50, pp.144-159 (2015).

Li H, Liu T, Wu X, et al., "An optimized VMD method and its applications in bearing fault diagnosis,". *Measurement*, Vol.166, No. 108185 (2020).

Liu X, Xie J, Luo Y, et al., "A novel power transformer fault diagnosis method based on data augmentation for KPCA and deep residual network". *Energy Reports*, Vol.9, pp. 620-627 (2023).

Navi M, Davoodi M R, Meskin N., "Sensor Fault Detection and Isolation of an Industrial Gas Turbine Using Partial Kernel PCA,". IFAC Papers OnLine, Vol.48(21), pp. 1389-1396 (2015).

Pacheco-Chérrez J, Fortoul-Díaz J A, Cortés-Santacruz F, et al., "Bearing fault detection with vibration and acoustic signals: Comparison among different machine leaning classification methods,". *Engineering Failure Analysis*, Vol.139, No.106515 (2022).

Pan Y, Hong R, Chen J, et al., "Incipient fault detection of wind turbine large-size slewing bearing based on circular domain,". *Measurement*, Vol.137, pp.130-142 (2019).

Ren H, Liu W, Shan M, et al., "A new wind turbine health condition monitoring method based on VMD-MPE and feature-based transfer learning,". *Measurement*, Vol.148, No. 106906 (2019).

Salameh J P, Cauet S, Etien E, et al., "Gearbox condition monitoring in wind turbines: A review,". *Mechanical Systems and Signal Processing*, Vol.111, pp.251-264 (2018).

Shanbr S, Elasha F, Elforjani M, et al., "Detection of natural crack in wind turbine gearbox,". *Renewable Energy*, Vol.118, pp.172-179 (2018).

Sharma, V., & Parey, A., "A review of gear fault diagnosis using various condition indicators,". *Procedia Engineering*, Vol.144, pp. 253-263 (2016).

Shen J, Xu F., "Method of fault feature selection and fusion based on poll mode and optimized weighted KPCA for bearings,". *Measurement*, Vol.194, No.110950 (2022).

Tong S, Huang Y, Jiang Y, et al., "The identification of gearbox vibration using the meshing impacts based demodulation technique,". *Journal of Sound and Vibration*, Vol.461, pp. 114879 (2019).

Touti W, Salah M, Bacha K, et al., "Gearbox mixed faults diagnosis under different working conditions based on meshing frequency modulation area,". *Applied Acoustics*, Vol.207, No.109336 (2023).

Wang S, Zhu R., "Research on dynamics and failure mechanism of herringbone planetary gearbox in wind turbine under gear surface pitting,". *Engineering Failure Analysis*, Vol.146, No. 107130 (2023).

Yaghoubi, V, Cheng, L , Paepegem, W. V. , & Kersemans, M., "A novel multi-classifier information fusion based on dempster–shafer theory: application to vibration-based fault detection,". *Structural Health Monitoring*, Vol.21(2), pp.596-612 (2022).

Yan, X, & Jia, M., "Application of csa-vmd and optimal scale morphological slice bispectrum in enhancing outer race fault detection of rolling element bearings,". *Mechanical Systems and Signal Processing*, Vol.122, pp. 56-86 (2019).

Yang X, Tofighi-Niaki E, Zuo M J, et al., "Analysis of spur gearbox dynamics considering tooth lubrication and tooth crack severity progression,". *Tribology International*, Vol.178, No.108027 (2023).

Zhang X, Zhong J, Li W, et al., "Nonlinear dynamic analysis of high-speed gear pair with wear fault and tooth contact temperature for a wind turbine gearbox,". *Mechanism and Machine Theory*, Vol.173, No.104840 (2022).

Zheng L, Lu W, Zhou Q., "Weather image-based short-term dense wind speed forecast with a ConvLSTM-LSTM deep learning model,". *Building and Environment*, Vol.239, No. 110446 (2023)

Zhu Y, Zhu C, Song C, et al., "Improvement of reliability and wind power generation based on wind turbine real-time condition assessment,". *International Journal of Electrical Power & Energy Systems*, Vol.113, pp.344-354 (2019).

基於多源特徵融合的風電機組齒輪箱早期故障診斷

李旺 何建軍 王佳煒 唐志偉 揭軍
羅光武 劉琰
國家能源集團湖南電力有限公司
長沙理工大學能源與動力工程學院
龍源江永風力發電有限公司

摘 要

作為風力渦輪機的主要運動部件，齒輪箱的故障率很高，對設備的危害特別大。目前使用的典型振動檢測技術對早期齒輪箱問題信號的診斷效果較差。考慮到這一點，本研究基於 KPCA-VMD 方法，提出了一種風力渦輪機齒輪箱早期故障監測和多維特徵評估方法，用於分析風力渦輪機齒輪箱不明顯的早期故障信號。首先，對預處理的數據集進行特徵提取，通過 KPCA 方法對齒輪箱特徵數據進行降維和重構，使用兩個統計量 T^2 和 SPE 監測齒輪箱狀態，並通過 VMD 分析監測到的異常信號。實驗數據表明，該方法可以有效診斷齒輪的早期故障特徵頻率。