Shi P, Yang W.

Precise Feature Extraction from Wind Turbine Condition Monitoring Signals by using Optimized Variational Mode Decomposition.

***IET Renewable Power Generation* 2016**

DOI: http://dx.doi.org/10.1049/iet-rpg.2016.0716

Precise Feature Extraction from Wind Turbine Condition Monitoring Signals by using Optimized Variational Mode Decomposition

Pu Shi[1*], Wenxian Yang[1]

[1] School of Marine Science and Technology, Newcastle University, Newcastle upon Tyne, UK, NE1 7RU

*Pu.Shi@newcastle.ac.uk

Abstract: Reliable condition monitoring (CM) highly relies on the correct extraction of fault-related features from CM signals. This equally applies to the CM of wind turbines (WTs). But influenced by slowly rotating speeds and constantly varying loading, extracting fault characteristics from lengthy, nonlinear, non-stationary WT CM signals is extremely difficult, which makes WT CM one of the most challenge tasks in wind power asset management despites that lots of efforts have been spent. Attributed to the superiorities to Empirical Mode Decomposition (EMD) and its extension form Hilbert-Huang Transform (HHT) in dealing with nonlinear, non-stationary CM signals, the recently developed Variational Mode Decomposition (VMD) casts a glimmer of light for the solution for this issue. However, the original proposed VMD adopts default values for both number of modes and filter frequency bandwidth. It is not adaptive to the signal being inspected. As a consequence, it would lead to inaccurate feature extraction thus unreliable WT CM result sometimes. For this reason, a precise feature extraction method based on optimised VMD is investigated in this paper. The experiments have shown that thanks to the use of the proposed optimisation strategies, the fault-related features buried in WT CM signals have been extracted out successfully.

Keywords: wind turbine, condition monitoring, variational mode decomposition, fault diagnosis

1. Introduction

As faults occurring in a Wind turbine (WT) may degenerate power generation performance, increase downtime, rise maintenance costs, and even lead to catastrophic failure in worst cases [1], WT condition monitoring (CM) that aims to detect these faults at an early stage has been regarded as one of the most effective measures to lower the Cost of Energy (COE) of wind power [2-6]. This is why WT CM is attracting increasing interest today from both academic and the industrial communities [7-9].

WT CM can be implemented in many ways [5]. But to date, the most proven approach is still based on signal processing because in contract to those mathematical models that are used to describing the dynamic behaviour of a WT, the CM signals collected directly from an operating WT can give more trustful, reliable and accurate description of the actual health condition of the WT being inspected. However, WTs are subjected to constantly varying loading conditions; moreover in order to maximize power generation at wide range of wind speeds, WTs are controlled by using a variable speed strategy. For these reasons, the CM signals collected from WTs are always nonlinear and non-stationary in both time and frequency domains. In contract to those CM signals collected from fixed-speed steam turbines operating in nuclear or coal fire power plants, the interpretation of the CM signals from WTs is more difficult. Therefore, the detection of incipient WT drive train fault is nearly impossible if without the aid of an appropriate signal processing tool [10-14].

In order to interpret non-stationary CM signals, many signal processing methods have been developed, such as short-time Fourier transform (STFT), Wigner-Ville distribution (WVD), wavelet transform, S-transform [15] and so on [16-18]. However, as the STFT is originated from Fourier transform, it is limited to Heisenberg's uncertainty principle as well; the WVD shows disadvantage due to the cross-term interference; the multi-resolution strategy of wavelet transform does benefit feature extraction from non-stationary signals. But in essence, it is a linear method based on fixed mother wavelet function, which is not adaptive to the signal being inspected thus still cannot fully guarantee the precise extraction of the local characteristics of the signal; S-transform takes the advantages of both the STFT and wavelet transform. It is more efficient in calculation than wavelet transform and more powerful in feature extraction than the STFT [15]. However, it also inherits the disadvantages of both pioneer methods, i.e. it is not locally adaptive as well. Hence, the locally adaptive feature of the empirical mode decomposition (EMD) and its extended form local mean decomposition (LMD) becomes an attractive solution for processing nonlinear and non-stationary CM signals [19]. Moreover,

they both have been applied to WT CM. However, as the EMD and LMD decompose the signal via recursive calculations [20], which show the following disadvantages:

- *the recursive calculations decreases the efficiency of calculation.* As a consequence, the EMD and LMD could face difficulty in dealing with lengthy CM signals, e.g. the WT CM signals;
- *the estimation of the upper and lower envelops is one of the key steps in EMD and LMD calculations.* Therefore, the error developed in envelop estimation will spread in the recursive decomposition results;
- *as mentioned in* [21], *the intrinsic mode functions (IMFs) resulted by the EMD are not band-limited functions*. Thus, the EMD is always accompanied by mode mixing issues [22], which significantly affects the accuracy of signal processing. To date, a fully successful solution for such an issue has not been achieved despite the lots of efforts, such as the development of the ensemble EMD (EEMD) and the ensemble LMD (ELMD) [23-26]. Practice has shown that the mixing modes existing in the IMFs prevent the correct presentation of the signal's time-frequency features by the approach of Hilbert transform.

To overcome these issues, an alternative non-recursive signal decomposition method, namely variational mode decomposition (VMD), was proposed recently [27]. In contrast to the EMD, the VMD is implemented in frequency domain. In other words, the VMD decomposes a signal into a few principal modes, then these modes will be constantly updated with the aid of Wiener filtering technique. The central frequency of each mode will be gradually demodulated to the corresponding baseband. Since the VMD is based on non-recursive calculations, it is more efficient in computing algorithm than the EMD and LMD. Moreover, the application of Wiener filtering technique enables the VMD to be very much robust against the negative influences of background noise on the accuracy of signal processing [27]. In addition, attributed to the application of Wiener filters, the VMD resultant modes are narrow-banded functions, which not only skilfully mitigates the mode mixing issues existing in the EMD, but also benefits the accurate extraction of the time-frequency features by the approach of Hilbert transform.

Since the VMD method was proposed, many researchers have explored its potential application in CM signals feature extraction. For example, a novel method of rubbing fault diagnosis based on the VMD was proposed in [28] and the equivalent filtering characteristics

of the VMD were also investigated in this paper. A novel fault diagnosis method based on the VMD and multi-kernel support vector machine is proposed in [29]. In [30], the performance comparison of the VMD and empirical wavelet transform (EWT) for the classification of power quality by using support vector machine was discussed. The results show that the VMD outperforms over EWT in feature extraction. To alleviate the negative influence of wind power generation, the VMD based novel model is proposed to improve the predication performance for the short-term wind power forecasting [31].

In [32], the superiorities of the VMD over the EMD was verified and the potential application of the VMD to the WT CM was also investigated. Moreover, the authors have also pointed out the proposed VMD shows the following two defects, which limits its feature extraction capability to a large extent:

- ***the number of modes K***. Different from the EMD, the number of modes K resulted by the VMD is controllable. It should be defined in advance before starting VMD calculation. In theory, the value of K is dependent on the number of frequency components n included in the signal. Once the value of K is defined inappropriately, not only unnecessary calculation would be increases, but also the accuracy of signal decomposition could be influenced. Therefore, it is very necessary to optimize the value of K to ensure the accuracy of VMD and its efficient calculation. However, the conventional VMD fails to perform such an optimization. For example, a default value $K = 3$ is applied to processing all signals in [27];
- ***bandwidth control parameter a***. Wiener filtering is the key technique adopted by the VMD. In conventional VMD, much effort has been spent on locate the central frequencies of the modes and the filters. An equivalent bandwidth is used to construct all filters with different central frequencies, which is not right particularly when processing a narrow-band mode function with large sideband noise. So, the frequency bandwidth of the filter should be adaptive to the mode function being inspected. Therefore, the bandwidth control parameter a in the VMD calculation should be optimized as well to fully reduce the influence of background noise on the premise of correct feature extraction. However, the conventional VMD does not take any consideration of this influence and uses a default value of a in all calculations.

It has no doubt that the utilization of inappropriate u_k and a will lead to the degeneration of the performance of the VMD and thus discount its added value in the application of WT CM. As an extension of the work reported in [32], the optimization of the VMD is studied in this

paper in order to overcome the aforementioned issues of the VMD. Moreover, the benefit of performing such an optimization is experimentally demonstrated by applying the optimized VMD to interpreting the CM signals collected from a WT bearing. More details about this work is depicted in the following sections.

2. Fundamental theory of the VMD

The VMD aims to decompose the signal into different discrete modes. These modes have specific sparsity properties, i.e. it is assumed that each mode is compactly distributed around its central frequency. The procedures to perform the VMD process are described as follows [27]:

a) For a given signal, define the number of modes K and mode bandwidth control parameter a, and initialize the modes $u_k(t)$ $(k = 1,2,3, \dots K)$;

b) For each mode $u_k(t)$, compute the associated analytic signal by means of Hilbert transform, i.e.

$$\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t) \tag{1}$$

c) Shift the frequency spectrum of each mode to the respective estimated central frequency

$$\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{-j\omega_k t} \tag{2}$$

d) The bandwidth is estimated through the squared L2-norm. The resulting constrained variational problem is given by the following

$$\min_{\{u_k\},\{w_k\}}\left\{\sum_k\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{-j\omega_k t}\right\|^2\right\} \quad s.t.\sum_k u_k = f \tag{3}$$

where, $\{u_k\} \coloneqq \{u_1,\dots,u_k\}$, and $\{w_k\} \coloneqq \{w_1,\dots,w_k\}$ represent the identified set of modes and their central frequencies.

A quadratic penalty term α and a Lagrangian multiplier factor $\lambda(t)$ are introduced in order to render the unconstrained problem. The augmented Lagrange expression is as follows

$$\begin{aligned} L(\{u_k\},\{w_k\},\lambda) = \alpha\sum_k\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{-j\omega_k t}\right\|_2^2 \\ +\left\|f(t)-\sum_k u_k(t)\right\|_2^2+\left\langle\lambda(t),f(t)-\sum u_k(t)\right\rangle \end{aligned} \tag{4}$$

An alternate direction multipliers method is applied to find the saddle point of the augmented Lagrange expression.

$$u_k^{n+1} = \arg\min_{u_k \in X} \left\{ \begin{array}{l} \alpha \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \\ + \left\| f(t) - \sum_i u_i(t) + \frac{\lambda(t)}{2} \right\|_2^2 \end{array} \right\} \tag{5}$$

where, w_k and $\sum_i u_i(t)$ correspond to w_k^{n+1} and $\sum_{i \neq k} u_i(t)^{n+1}$, respectively. Convert (5) into spectral domain and apply the Parseval/Plancherel Fourier isometry.

$$\hat{u}_k^{n+1} = \arg\min_{\hat{u}_k, u_k \in X} \left\{ \begin{array}{l} \alpha \left\| j\omega \left[(1 + \mathrm{sgn}(\omega + \omega_k)) \hat{u}_k(\omega + \omega_k) \right] \right\|_2^2 \\ + \left\| \hat{f}(\omega) - \sum_i \hat{u}_i(\omega) + \frac{\lambda(\omega)}{2} \right\|_2^2 \end{array} \right\} \tag{6}$$

With manipulation ω with $\omega - \omega_k$,

$$\omega_k^{n+1} = \arg\min_{\omega_k} \left\{ \int_0^\infty (\omega - \omega_k)^2 \left| \hat{u}_k(\omega) \right|^2 d\omega \right\} \tag{7}$$

By exploiting the Hermitian symmetry, (7) can be further inferred

$$\hat{u}_k^{n+1} = \arg\min_{\hat{u}_k, u_k \in X} \left\{ \begin{array}{l} \int_0^\infty 4\alpha (\omega - \omega_k)^2 \left| \hat{u}_k(\omega) \right|^2 \\ +2 \left| \hat{f}(\omega) - \sum_i \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2} \right| d\omega \end{array} \right\} \tag{8}$$

The solution of this quadratic optimization problem is as follows

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha (\omega - \omega_k)^2} \tag{9}$$

Then, the central frequency, ω_k , can be converted to the frequency domain

$$\omega_k^{n+1} = \arg\min_{\omega_k} \left\{ \int_0^\infty (\omega - \omega_k)^2 \left| \hat{u}_k(\omega) \right|^2 d\omega \right\} \tag{10}$$

The quadratic equation (10) is easily solved as

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega \left| \hat{u}_k(\omega) \right|^2 d\omega}{\int_0^\infty \left| \hat{u}_k(\omega) \right|^2 d\omega} \tag{11}$$

From the theory depicted above, it can be inferred that the original proposed VMD shows the following limitations in signal feature extraction. Firstly, the number of decomposition

number K needs to be defined in advance without any pre-knowledge about the signal being inspected. Thus, it is difficult to guarantee that the appropriateness of setting of the value of K. This would affect the accuracy of feature extraction; Secondly, mode frequency bandwidth control parameter *a* should be carefully selected in order to suppress noise interference as much as possible. In principle, it is adaptive to the mode function being processed. However, a fixed default value of *a* is adopted by the conventional VMD. Apparently, both these two limitations need to improve to reduce the error of the VMD in feature extraction, which motivates the following research.

3. Parameter Optimisation of the VMD

3.1. Optimum number of modes

It is well known that WT CM signals are often contaminated by background noise and electromagnetic interferences, which smear the fault-related features contained in the signals and as a consequence, lower the reliability of CM conclusion. Thus, take measures to mitigate these undesired influences is essential. The optimization of the number of modes depicted in this section is just an effort for reaching such a purpose. Considering the background noise and electromagnetic interferences in frequency spectrum are often smaller in magnitude in contrast to valid frequency components, the following spectrum-envelope based method is developed to identify the valid frequency components contained in the signal, and according to which to furthermore determine the optimum number of modes of the VMD. The methods can be implemented by following the steps described below.

Step 1: Implement the Fourier transform of the CM signal *x*

$$F(\mathrm{s}) = \mathrm{FFT}(\mathrm{x}) \tag{12}$$

$$A_{ax} = |F(s_i)| \quad (i = 1,2,\cdots,l) \tag{13}$$

where, l is an integer equal to half number of data x.

Step 2: Detect the local maximum and local minimum magnitudes of A_{ax}.

Step 3: Generate the envelope curve of the filtered spectrum by using the cubic spline interpolation method. The relevant calculation can be easily implemented by using an available Matlab function namely ‘spline.m’.

Step 4: Calculate the following threshold, based on which to filter out noise and interference components

$$Threshold = A_l + r(A_h - A_l) \tag{14}$$

where, A_h and A_l are respectively the maximum and minimum magnitudes that can be observed from the spectrum, r is a ratio used for controlling the level of the threshold, i.e. the larger the value of r, the higher the level of the threshold tends to be. In this paper, r was taken to be 0.2 based on the assumption is that the signal-to-noise ratio of WT CM signals is usually higher than 85 dB.

Step 5: Filter out those frequency components that have smaller magnitude than the threshold value through setting them to be zero.

Step 6: Identify the optimum number of modes from the resultant envelope curve.

To ease understanding, a bearing CM signal is taken as an example to demonstrate the aforementioned algorithm. The time waveform of the signal is shown in Fig.1.

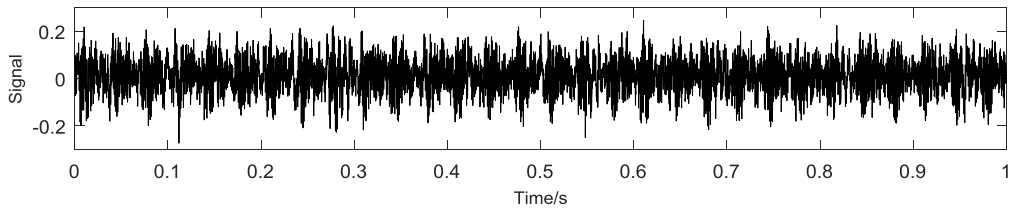
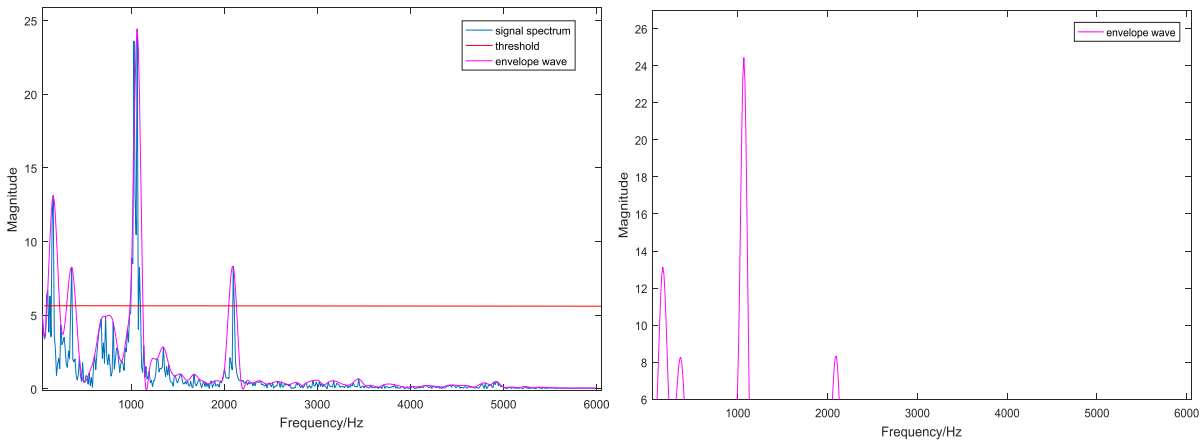


Fig.1. An example of bearing condition monitoring signal

Calculate the frequency spectrum and the associated envelope curve of the signal, see Fig.2a. Apply (14) to calculating the value of the threshold from the associated envelope curve and obtain *Threshold* = 6. Then, apply this threshold to Fig.2a, the trimmed envelope curve is shown in Fig.2b.



(a) envelope curve (b) filtered envelope curve

Fig.2. Mode number detection illustration for bearing under normal condition

From Fig.2b, it is clearly seen that there are only 4 valid frequency components remains in the spectrum, thus the right number of modes for the VMD calculation should be set as 4.

The mode number detection result obtained above can be readily validated with the aid of correlation analysis between of the raw signal and the purified signals constructed by using the VMD modes. It can be imagined that the more number of modes are used to reconstruct the signal, the reconstructed signal will be more correlated to the raw signal. Thus, with the increase of the number of modes for purified signal construction, the resultant correlation coefficient of the purified signal and the raw signal will increase gradually until reaching a saturated value. Then, the number of modes that corresponds to the saturated value of the correlation coefficient will indicate the optimum value of the mode number K. A smaller mode number will lead to the phenomenon of mode mixing in VMD result, while a larger mode number will not only lead to unnecessary calculations thus decrease computational efficiency of the VMD, but also result in the presence of a number of modes with redundant information. To demonstrate such a point of view and verify the correctness of the aforementioned method for mode number optimization, the CM signal shown in Fig.1 is dealt with by using the VMD with the setting of a significant number of modes 50. Then, calculate the signal correlation coefficient ratio of the modes by using an equation

$$C(n) = corr\left(\sum_{k=1}^{n} m_k, x\right) \tag{15}$$

where, n increases from 1 to 50, $corr(m_k, x)$ is the correlation coefficient between the kth VMD mode and the input signal x.

The variation tendency of the resultant correlation coefficient ratio $C(n)$ against the number of modes is shown in Fig.3.

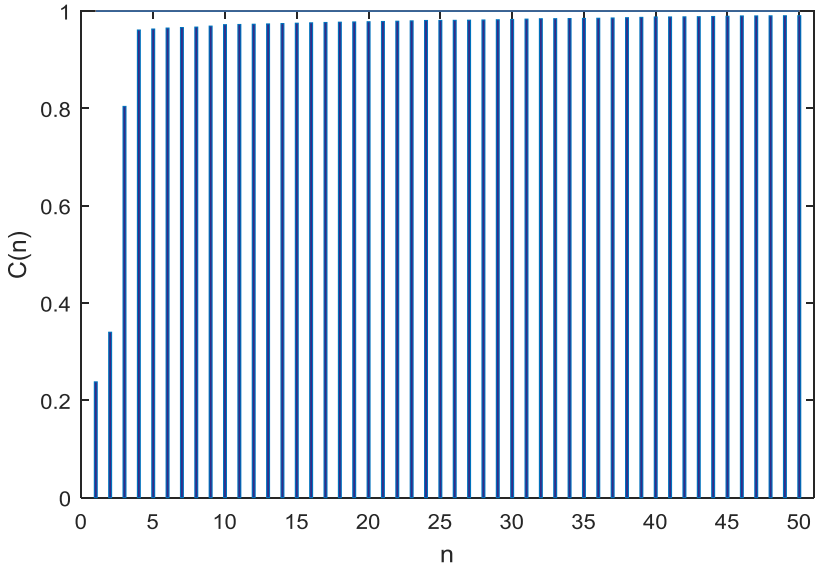


Fig. 3. Variation tendency of the correlation coefficient ratio

From Fig.3, it is found that the resultant correlation coefficient increases gradually with the increase of the number of modes, and finally reaches a saturated value when the number of mode is 4. This is in agreement with the aforementioned detection result very well and indicates once again that the first 4 VMD modes should have contained almost all valid information in the raw CM signal. Through this example, it can be said that method proposed in this section is valid for detecting the optimum number of VMD modes.

3.2. Optimum frequency bandwidths of Wiener filters

According to [27] and also the algorithm depicted in Section 2, the frequency bandwidth control parameter α will significantly influence the VMD results. As Fig.4 shows, the smaller the parameter α is, the wider the frequency bandwidth of the filter tends to be, and vice versa. When the bandwidth of the filter is wide, more information, such as background noise and interference frequency items, will be included in the filtering result. This will decrease the signal-to-noise ratio of the filtering result and increase the difficulty of signal interpretation. However, the frequency bandwidth control parameter α cannot be too big as a narrow frequency bandwidth of the filter will raise the risk of that important parts of the signal are either discarded as noise when the number of modes is too few or shared by adjacent modes that have same centre frequency when the number of modes is too many. In summary, decreasing the bandwidth by increasing α comes at the risk of not properly capturing the correct mode, while α too low includes more noise in the estimated mode. Thus, the parameter α, i.e. the frequency bandwidth of Wiener filter, should be optimized in the calculation of VMD to ensure the correctness of the VMD results.

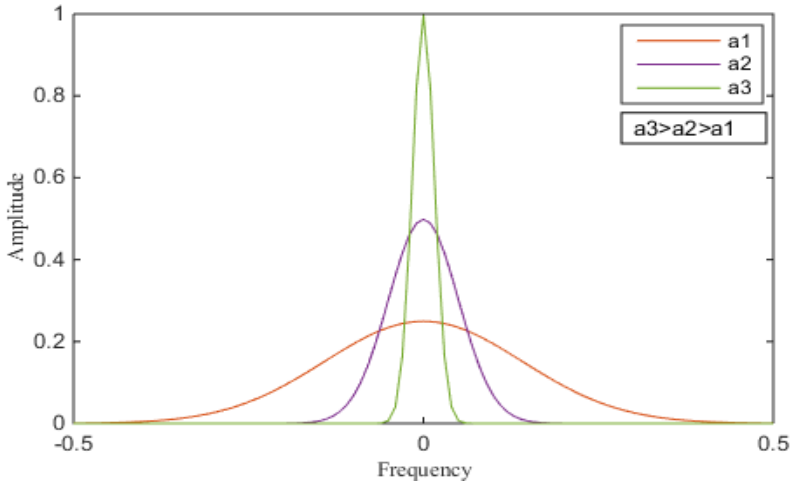


Fig. 4. Influence of parameter α on the frequency bandwidth of Wiener filter

To achieve such a purpose, an algorithm is developed to detect the optimum value of parameter a. The algorithm is depicted as follows.

Step 1: For an input signal x, detect the optimum number of modes by using the algorithm described in Section 3.1, and then set n to be the optimum number of modes;

Step 2: Define an initial value of parameter $a = 200$ and decompose the signal x by using the VMD, i.e.

$$x = \sum_{k=1}^{n} m_k \tag{16}$$

where m_k denotes the k-th VMD mode.

Step 3: Calculate the total contribution of all centre frequency components energy of centre frequency (EOC) in VMD results, i.e.

$$EOC = \sum_{k=1}^{n} A_k \tag{17}$$

As defined in (15), A_k is the maximum magnitude observed from the frequency spectrum of the kth VMD mode;

Step 4: Calculate the total energy of all modes (EOM) from their frequency spectra

$$EOM = \sum_{k=1}^{n} \sum_{j=1}^{N} M_{k,j} \tag{18}$$

where, $M_{k,j}$ represents the magnitude of the j-th frequency component in the spectrum of the k-th VMD mode. N is equal to the half number of data contained in the mode;

Step 5: Calculate and record the value of signal energy distribution ratio (SEDR)

$$SEDR = \frac{EOC}{EOM} \tag{19}$$

Step 6: Increase the value of a and repeat Steps 2 to 5 until a reaches 10000;

Step 7: Plot the variation tendency of the SEDR against the value of a.

For the bearing CM signal shown in Fig.1, the resultant tendency of SEDR versus a is shown Fig.5.

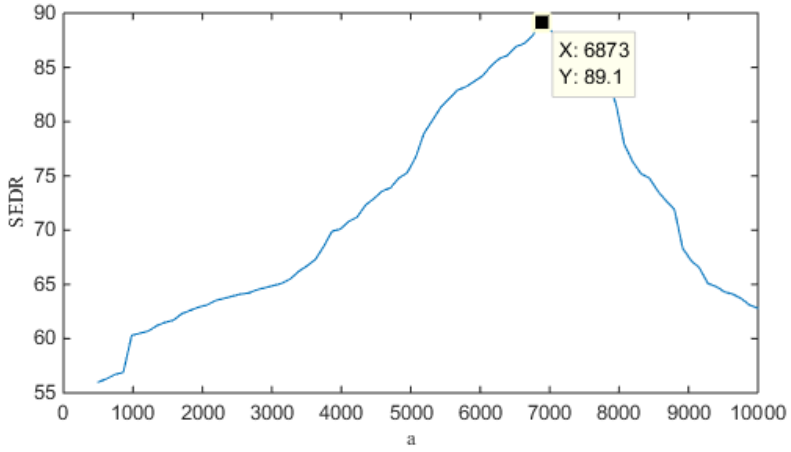


Fig.5. SEDR versus a

From Fig.5, it is interestingly found that at the beginning, the SEDR increases gradually with the increasing value of a until it reaches the maximum value when a = 6873, since which the SEDR turns to decrease with the increase of a. This implies that filter with parameter a = 6873 has the best signal-to-noise ratio.

Based on the above description, the VMD with the optimized parameters of K and a can be implemented by using the following algorithm:

1) Determine the optimum number of modes K by using the method depicted in Section 3.1;
2) Determine the optimum mode bandwidth control parameter a by using the method proposed in Section 3.2;
3) Initialize $\{\hat{u}_k^1\}$, $\{\omega_k^1\}$, $\{\hat{\lambda}^1\}$, and $n = 0$;
4) Update n with $n + 1$, and iterate the following loop calculations until $k = K$, i.e.

   $For\ k = 1{:}K$

   $Update\ \hat{u}_k\ for\ all\ \omega \geq 0\ with$

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i<k} \hat{u}_i^{n+1}(\omega) - \sum_{i>k} \hat{u}_i^n(\omega) + \frac{\hat{\lambda}^n(\omega)}{2}}{1 + 2a(\omega - \omega_k^n)^2}$$

   $Update\ \omega_k\ with$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2\, d\omega}{\int_0^\infty \left|\hat{u}_k^{n+1}(\omega)\right|^2 d\omega}$$

   $End\ for$

   Then, dural ascent for all $\omega \geq 0$

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \tau(\hat{f}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega))$$

Until the following convergence condition is satisfied

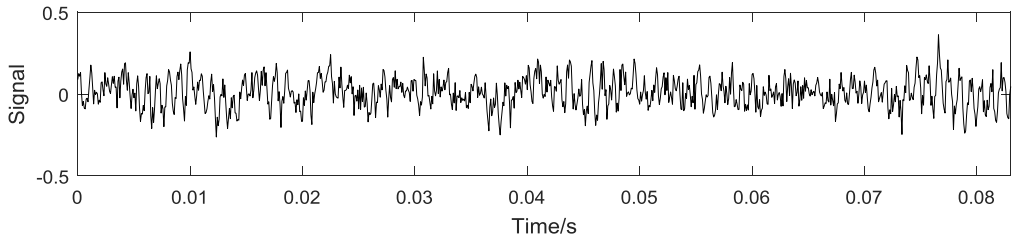
$$\sum_{k=1}^{K} \frac{\left\|\hat{u}_k^{n+1}-\hat{u}_k^n\right\|_2^2}{\left\|\hat{u}_k^n\right\|_2^2} < \varepsilon$$

4. Feature extraction by using the VMD with the optimum parameter values

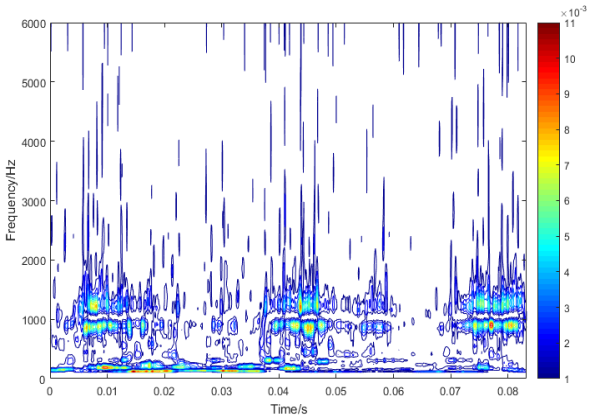
In this section, the contribution of the aforementioned algorithms to improving the quality of VMD results is experimentally demonstrated with the aid of the CM signals collected from a test rig that was designed dedicatedly for testing the reliability of WT bearings. On the test rig, the rotating speed of the bearing is controlled by a DC motor, which can rotate at variable speeds to simulate the operating conditions of the bearing at different wind speeds.

4.1. VMD based CM application for normal bearing

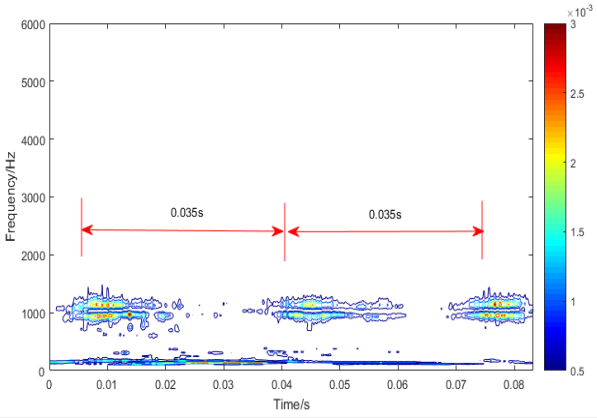
First of all, the bearing vibration signal was collected by using a sampling frequency of 12 kHz when the bearing was health and without any defect. The time waveform of the signal is shown in Fig.6a. To facilitate comparison, both the original proposed VMD and the VMD with optimized parameters were applied to dealing with the signal. The corresponding results are shown in Figs.6b and 6c.



(a) Time waveform of the signal



(b) VMD result when $n = 3$, $a = 500$

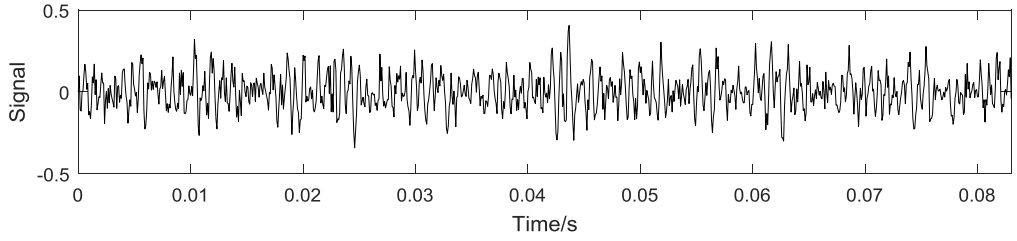


(c) VMD result when $n = 4$, $a = 6873$

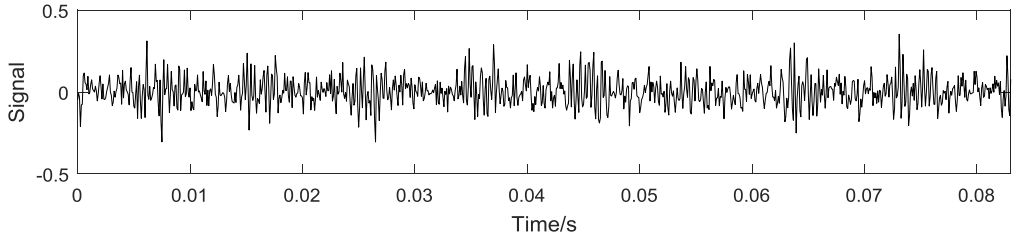
Fig.6 Feature extraction results when the bearing is normal

From Fig.6a, it is seen that the vibration signal collected even when the bearing is under normal condition is still very noisy, i.e. almost all bearing rotational features are hidden by the noise. Moreover, the noise often has very wide frequency bandwidth. As a consequence, it could mask the signal features in a wide range of frequencies. The comparison of Figs.6b and 6c shows that after using the optimum values of the parameters K and a, the noise contained in the vibration signal has been cancelled sufficiently. In other words, the energy of the signal spreads over a wide range of frequencies from 0 to 6 kHz in the results decomposed by using the conventional VMD. Moreover, the rotational feature of the bearing is seriously smeared by the noise. By contrast, after using the improved VMD with the optimum values of the parameters K and a, most noise is successfully cancelled and thanks to which the rotational feature of the bearing is clearly presented in the time-frequency map of the signal. In Fig6c, the time interval 0.035s indicates the bearing rotating speed of 1750 rev/min.

To demonstrate the significant benefit of using the optimized VMD in extracting fault-related features from WT CM signals, the proposed parameter optimization algorithm was applied to processing the vibration signals collected respectively when defects occurred on the inner race and outer race of the bearing. The raw signals collected in both scenarios are showing in Fig.7.



(a) Inner race fault

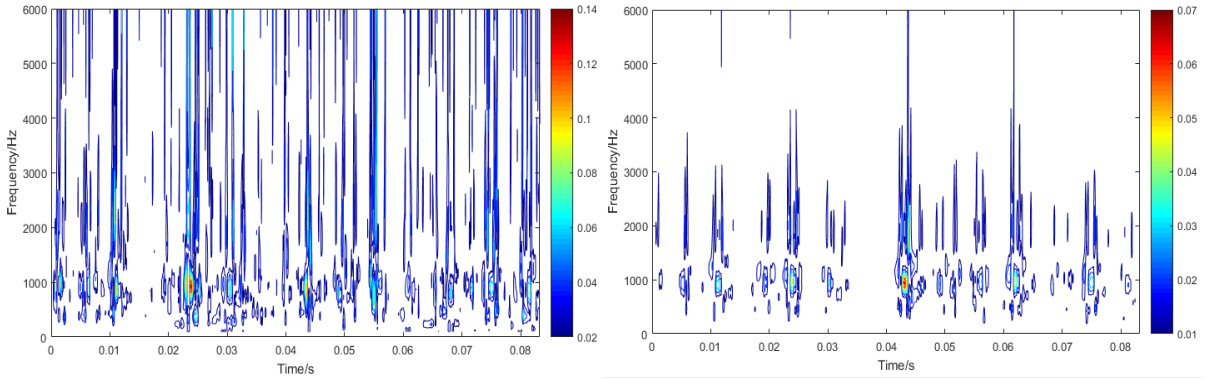


(b) Outer race fault

Fig.7 Vibration signals collected when the bearing was faulty

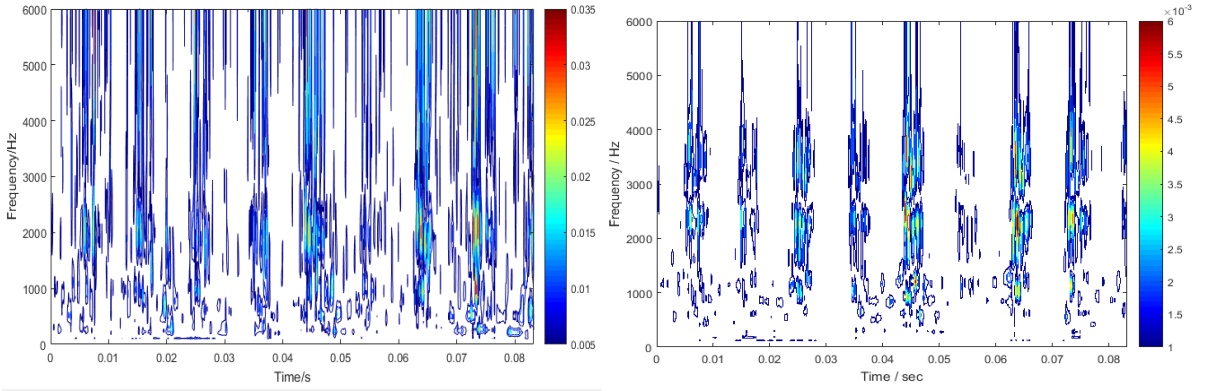
As can be imagined, impact features are present in the vibration signals once a defect is developed on the surface of either inner race or outer race of the bearing. The time interval between the impact features discloses the characteristic frequency of the fault and thus becomes

an important clue in implementing fault diagnosis. For example, at the same bearing rotational speed, the outer race defect will result in a larger time interval than the inner race defect does. However, the practice, as seen in Fig.7, has shown that the bearing vibration signals are often contaminated by strong background noise due to various reasons (e.g. poor lubrication). Thus, how to accurately identify the time interval between impact features from noisy bearing vibration signals is a challenge. To address this issue, the optimized VMD was adopted for processing the signals in Fig.7 and the corresponding results are shown in Figs.8 and 9. Where the results obtained by using the conventional VMD are also shown for comparison.



(a) VMD result when $n = 3$, $a = 2000$ (b) VMD result when $n = 6$, $a = 7929$

Fig.8 VMD for the signal shown in Fig.7a



(a) VMD result when $n = 3$, $a = 2000$ (b) VMD result when $n = 7$, $a = 4036$

Fig.9 VMD for the signal shown in Fig.7b

From Figs.8a and 9a, it is found that when the default values of the mode number and bandwidth control parameter were adopted, the fault-rated impact features can be perceived from the time-frequency maps of the signals. However, considerable background noise is also present in the figures, which smears the fault features and consequently leads to the risk of either false alarm or the failure of fault detection. By contrast, Figs.8b and 9b disclose that after

optimizing the VMD parameters, the noise contained in the signals has been successfully filtered out by the VMD and thanks to this, the impact features contained in the signals are clearly presented in the time-frequency maps. Moreover, it can be readily measured that the impact features in Fig.8b show a time interval of about 0.006s, which corresponds to the characteristic frequency 157Hz of an inner race fault in the bearing; while the impact features in Fig.9b shows a time interval of about 0.009s, which corresponds to the characteristic frequency 104Hz of an outer race fault in the bearing.

5. Conclusions

The VMD has been proved superior to the EMD in both signal decomposition and feature extraction. However, the accuracy of the VMD results can be significantly affected by the number of mode K and mode frequency bandwidth control parameter α that are adopted in VMD calculations. In order to achieve an accurate VMD and thus explore a valid signal processing tool for WT CM, a parameter optimization algorithm was developed in this paper. The effectiveness of the proposed algorithm in processing WT CM signals has been experimentally verified. The experiments have shown that:

- An inappropriate value of the number of modes K not only decreases the algorithm efficiency of the VMD due to unnecessary calculation, but also lowers the accuracy of signal decomposition results; An inappropriate value of the mode bandwidth control parameter α will significantly discount the noise cancellation capability of the VMD and consequently reserve considerable noise in the VMD results, which will significantly increase the difficulty in signal interpretation;.
- With the aid of the proposed optimization algorithm, the accuracy of the VMD has been significantly improved, i.e. the impact features contained in the bearing vibration signals have been explicitly extracted out despite the considerable noise in the signals and the lack of any pre-knowledge about the CM signals.

Acknowledgement

The work reported in this paper was supported by China Natural Science Foundation with Reference Numbers of 11632011 and 11472103.

Reference

[1] Z. Hameed, Y. S. Hong, Y. M. Cho, S. H. Ahn and C. K. Song, "Condition monitoring and fault detection of wind turbines and related algorithms: A review," *Renewable and Sustainable Energy Reviews, 2009,* **13**, (1), pp. 1-39.

[2] M. Zaggout, P. Tavner, C. Crabtree and L. Ran, "Detection of rotor electrical asymmetry in wind turbine doubly-fed induction generators," *IET Renewable Power Generation, 2014,* **8**, (8), pp. 878-886.

[3] P. Tavner, D. Zappal and S. Sheng, "Side-band algorithm for automatic wind turbine gearbox fault detection and diagnosis," *IET Renewable Power Generation, 2014,* **8**, (4), pp. 380-389.

[4] W. Yang, C. Little, P. Tavner and R. Court, "Data-driven technique for interpreting wind turbine condition monitoring signals," *IET Renewable Power Generation, 2014,* **8**, (2), pp. 151-159.

[5] J. P. Barton and S. J. Watson, "Analysis of electrical power data for condition monitoring of a small wind turbine," *IET Renewable Power Generation, 2013,* **7**, (4), pp. 341-349.

[6] W. Yang, J. P. Tavner, J. C. Crabtree, Y. Feng and Y. Qiu, "Wind turbine condition monitoring: technical and commercial challenges," *Wind Energy, 2014,* **17**, (5), pp. 673-693.

[7] W. Yang, P. Tavner, C. Crabtree and M. Wilkinson, "Cost-effective condition monitoring for wind turbines," *IEEE Transactions on Industrial Electronics, 2010,* **57**, (1), pp. 263-271.

[8] W. Yang, P. J. Tavner and W. Tian, "Wind Turbine Condition Monitoring based on an Improved Spline-Kernelled Chirplet Transform," *IEEE Transactions on Industrial Electronics, 2015*.

[9] Y. G. Lei, J. Lin, Z. He and M. J. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery," *Mechanical Systems and Signal Processing, 2013,* **35**, (1-2), pp. 108-126.

[10] Y. Wang and D. Infield, "Supervisory control and data acquisition data-based non-linear state estimation technique for wind turbine gearbox condition monitoring," *IET Renewable Power Generation, 2013,* **7**, (4), pp. 350-358.

[11] W. Yang, P. Tavner, C. Crabtree, Y. Feng and Y. Qiu, "Wind turbine condition monitoring: technical and commercial challenges," *Wind Energy, 2014,* **17**, (5), pp. 673-693.

[12] J. Hang, J. Zhang and M. Cheng, "Fault diagnosis of wind turbine based on multi-sensors information fusion technology," *IET Renewable Power Generation, 2014,* **8**, (3), pp. 289-298.

[13] Y. Feng, Y. Qiu, C. Crabtree, H. Long and P. Tavner, "Monitoring wind turbine gearboxes," *Wind Energy, 2013,* **16**, (5), pp. 728-740.

[14] P. Tchakoua, R. Wamkeue, M. Ouhrouche, F. Slaoui-Hasnaoui, T. A. Tameghe and G. Ekemb, "Wind Turbine Condition Monitoring: State-of-the-Art Review, New Trends, and Future Challenges," *Energies, 2014,* **7**, (4), pp. 2595-2630.

[15] Y. Wenxian, C. Little and R. Court, "S-Transform and its contribution to wind turbine condition monitoring," *Renewable Energy, 2014,* **62**, pp. 137-146.

[16] Y. T. Sheen, "An envelope analysis based on the resonance modes of the mechanical system for the bearing defect diagnosis," *Measurement, 2010,* **43**, (7), pp. 912-934.

[17] W. He, Z. N. Jiang and K. Feng, "Bearing fault detection based on optimal wavelet filter and sparse code shrinkage," *Measurement, 2009,* **42**, (7), pp. 1092-1102.

[18] X. Chiementin, F. Bolaers and J. P. Dron, "Early Detection of Fatigue Damage on Rolling Element Bearings Using Adapted Wavelet," *ASME Journal of Vibration and Acoustics, 2007,* **129**, (4), pp. 495-506.

[19] H. Li, Y. Zhang and H. Zheng, "Hilbert-Huang transform and marginal spectrum for detection and diagnosis of localized defects in roller bearings," *Journal of Mechanical Science and Technology, 2009,* **23**, (2), pp. 291-301.

[20] M. Van, H. J. Kang and K. S. Shin, "Rolling element bearing fault diagnosis based on non-local means de-noising and empirical mode decomposition," *IET Science, Measurement & Technology, 2014,* **8**, (6), pp. 571-578.

[21] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceeding of the royal society, 1998,* pp. 903-995.

[22] Y. Yang, J. S. Cheng and K. Zhang, "An ensemble local means decomposition method and its application to local rub-impact fault diagnosis of the rotor systems," *Measurement, 2012,* **45**, pp. 561-570.

[23] Z. H. Wu and N. E. Huang, "Ensemble empirical mode decomposition:a noise assisted data analysis method," *Advances in Adaptive Data Analysis, 2009,* **1**, (1), pp. 1-41.

[24] G. Rilling, P. Flandrin and P. Goncalvès, "On empirical mode decomposition and its algorithms," *IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing, NSIP-03,* 2003.

[25] R. Li and D. He, "Rotational machine health monitoring and fault detection using EMD-based acoustic emission feature quantification," *IEEE Transactions on Instrumentation and Measurement, 2012,* **61**, (4), pp. 990-1001.

[26] W. Yang, R. Court, P. Tavner and C. Crabtree, "Bivariate empirical mode decomposition and its contribution to wind turbine condition monitoring," *Journal of Sound and Vibration, 2011,* **330**, (15), pp. 3766-3782.

[27] K. Dragomiretskiy and D. Zosso, "Variational Mode Decomposition," *IEEE Transactions on Signal Processing, 2014,* **62**, (3), pp. 531-544.

[28] Y. X. Wang, R. Markert, J. Xiang and W. Zheng, "Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system," *Mechanical Systems and Signal Processing, 2015,* **60-61**, pp. 243-251.

[29] Z. Lv, B. Tang, Y. Zhou and C. Zhou, "A Novel Method for Mechanical Fault Diagnosis Based on Variational Mode Decomposition and Multikernel Support Vector Machine," *Shock and Vibration,* 2016.

[30] C. Aneesh, S. Kumar, P. M. Hisham and K. P. Soman, "Performance Comparison of Variational Mode Decomposition over Empirical Wavelet Transform for the Classification of Power Quality Disturbances Using Support Vector Machine," in *roceedings of the International Conference on Information and Communication*

Technologies, ICICT 2014, 3-5 December 2014, Bolgatty Palace & Island Resort, Kochi, India, 2014.

[31] Y. Zhang, K. Liu, L. Qin and X. An, "Deterministic and probabilistic interval prediction for short-term wind power generation based on variational mode decomposition and machine learning methods," *Energy Conversion and Management, 2016,* 112, pp. 208-219.

[32] W. Yang, Z. Peng, K. Wei, P. Shi and W. Tian, "Superiorities of variational mode decomposition over empirical mode decomposition particularly in time–frequency feature extraction and wind turbine condition monitoring," *IET Renewable Power Generation*, 2016.