Wavelet Energy Transmissibility Function and Its Application to Wind Turbine Bearing Condition Monitoring

Long Zhang, *Member, IEEE*, and Zi-Qiang Lang

***Abstract*—Condition or health monitoring techniques and methods have been widely used for engineering systems fault detection and diagnosis. However, there is a major challenge with monitoring the systems operating under time varying loadings especially when the system loads are unknown or hard to measure. To address this problem, a new concept, wavelet energy transmissibility function, is proposed in this paper. The main advantage of this new method is that it can remove the impact of varying loadings but it does not require any loading information. Further, the proposed method is robust to noise and is sensitive to system property changes. The effectiveness of the proposed method has been well demonstrated by a numerical example, the theoretical study, and the analysis of the field vibration data from bearings of operating wind turbines.**

***Index Terms*—Wavelet energy transmissibility functions (WETF), wind turbine, condition monitoring.**

I. Introduction

FREQUENCY domain methods have been widely used for the condition monitoring of engineering systems. The fundamental principal is that damaged systems can produce frequency characteristics that is different from those under normal conditions. For example, some faults can cause frequency shocks, spikes or sidebands. If the monitored systems are under steady or stationary operating conditions, Fast Fourier Transform (FFT), modulation sidebands, envelope analysis, cepstrum analysis, skewness and kurtosis have been successfully used in many applications [1]–[3]. However, if the monitored systems are under non-stationary operating conditions, such as in the case of wind turbines, the frequency information may vary with the system dynamic loadings, which is referred to as the variable loading problems [4]–[10]. Variable loadings can result in difficulties in using traditional frequency domain techniques. To address the dynamic loading problems, time-frequency methods, such as short-time FFT, Empirical model decomposition, Hilbert transform, and wavelet transform and their modifications, can be used in many cases including both online and realtime and offline applications [11]–[16]. These methods are also well studied and used in other fields, such as structure health monitoring [17], [18].

Alternatively, the frequency response function (FRF) is a promising technique to remove the dynamic loading effects. FRF is defined as the spectra ratio between the dynamic response and corresponding loading input [19]–[21]. Due to the ratio operation, the effect of the input loading can often be eliminated, producing FRF solely dependent on system physical properties. For the purpose of health monitoring, the estimated FRFs can be compared with their baseline values [19], [22]. If the changes are beyond a warning threshold, the inspected system or component may have some physical damage. According to the definition of FRF, the computation of the FRF needs to measure the system input and response simultaneously. However, the information of input loading is often not available. In the case, FRF can not be used for condition monitoring purpose.

As a similar concept to FRF, transmissibility function (TF) is defined as the ratio of two different FRFs and is also equal to the ratio of the spectra of two different responses. In other words, unlike FRF which requires both system input and outputs, TF only requires system outputs or measured responses. TF can be used to represent the system physical properties. For example, for a multi-degree of freedom (MDOF) system, TF is solely dependent on modal parameters including mass, stiffness and damping [22]–[25]. TF based condition monitoring has been widely used in many applications [21], [24], [26]–[28]. A good review on the transmissibility analysis can be found in [27].

In practical applications, FRF and TF are often estimated from measured data using Fourier transform [29]. Most recently, instead of Fourier transform, the wavelet transform is proposed to compute FRF [30] and experimental investigations on wavelet based FRF with validations to detect abrupt change in stiffness were carried out in [31]. The new wavelet FRF is the ratio of coefficients of wavelet transforms of system input and output. A similar type of concept employing the ratios of wavelet coefficients at different frequency bands were proposed in [32] to identify time-varying and nonlinear systems. Due to the relationship between FRF and TF, wavelet FRF can be naturally extended to wavelet TF (WTF) as the ratio between wavelet coefficients of two different system responses. However, it is

Manuscript received January 24, 2017; revised September 3, 2017 and December 16, 2017; accepted February 22, 2018. Date of publication March 29, 2018; date of current version September 18, 2018. Paper no. TSTE-00080-2017. (*Corresponding author: Long Zhang).*

L. Zhang is with the School of Electrical and Electronic Engineering, University of Manchester, Manchester M13 9PL, U.K. (e-mail: long.zhang@manchester.ac.uk).

Z. Lang is with the Department of Automatic Control of System Engineering, University of Sheffield, Sheffield S1 3JD, U.K. (e-mail: z.lang@sheffield.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSTE.2018.2816738

found in the present study that it is often difficult to produce an consistent WTF using real data. The main reason for this can be due to the variable or dynamic loading effect as the estimated WTF may have different values under different loading conditions. Another reason can be due to the noise effect which can easily corrupt useful information in data, especially data with small amplitudes.

In the present study, a new concept, wavelet energy transmissibility function (WETF), is proposed to overcome the effects of dynamic loadings and noise on the data analysis and analyzed in theory. The proposed method is evaluated using numerical examples and a large amount of field data from operating wind turbine bearings and promising results are achieved.

II. Wavelet Transmissibility Function (WTF)

When system responses are measured by multiple sensors simultaneously, the wavelet transmissibility can be calculated using the ratio of coefficients resulting from wavelet transforms of two different measurements. Therefore, the wavelet coefficients need to be first obtained. Suppose there are n measurements which are denoted as $[X_1, \ldots, X_i, \ldots, X_n]$ where $X_i = [x_i(1), \ldots, x_i(N)]^T$ with N being the data length. The wavelet transform of $X_i, i = 1, \ldots, n$, is given by

$$W_i(a,b) = a^{-\frac{1}{2}} \int x_i(t) \varphi \left(\frac{t-b}{a} \right) dt \tag{1}$$

where φ is a wavelet function and a, b are dilation and translation parameters, respectively. The integral operation is computationally expensive, therefore a direct method is not preferable in practice. To improve the computational efficiency, the multiple resolution analysis (MRA) method was proposed in [33], which makes wavelet analysis be widely used in all aspects of engineering and science fields due to its excellent computational efficiency. The MRA uses multiple levels decomposition to compute coarse-to-fine frequency resolutions. In each level of decomposition, MRA first involves computing the convolutions with high pass filter g and low pass filter h where the filters are determined by the wavelet orthonormal basis, and then adopts downsamplings for both filtered results. For level one decomposition, the original signal X_i is decomposed to a detailed part D_1^i and an approximation part A_1^i. More specifically, the decomposition process can be denoted as

$$D_1^i = (X_i * g) \downarrow 2 \tag{2}$$

$$A_1^i = (X_i * h) \downarrow 2 \tag{3}$$

where

$$X_i * g = \left\{ \sum_k x_i(k) g(q-k), q = 1, \ldots, N \right\} \tag{4}$$

$$X_i * h = \left\{ \sum_k x_i(k) h(q-k), q = 1, \ldots, N \right\} \tag{5}$$

and $\downarrow 2$ indicates the downsampling by 2. To make it clear, the level 1 decomposition is illustrated in Fig. 1. The approximation part A_1^i can be further decomposed using the same procedure.

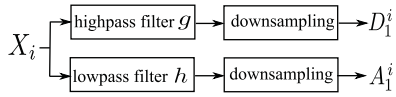

Fig. 1. Wavelet level 1 decomposition.

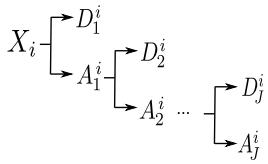

Fig. 2. J level wavelet decomposition.

The second decomposition is called level 2 decomposition. The decomposition continues until a pre-set level is satisfied. For the wavelet transform with J level decomposition, it can be described in Fig. 2 and wavelet coefficients are given by

$$\begin{aligned} W_i &= \{A_J^i, D_J^i, D_{J-1}^i, \ldots, D_1^i\} \\ &= \{\nu_i(r), r = 1, \ldots, R\} \end{aligned} \tag{6}$$

The TF is the ratio between two different measurements. Let ith and jth denote the two different measurement indexes, where $i = 1, \ldots, n-1, j = i+1, \ldots, n$, and the transmissibility function between two measurements can be written as

$$T_{ij} = [t_{ij}(1), \ldots, t_{ij}(r), \ldots, t_{ij}(R)] \tag{7}$$

where

$$t_{ij}(r) = \frac{\nu_i(r)}{\nu_j(r)} \tag{8}$$

The total number of such transmissibility functions is $L = (n-1)n/2$. These functions will be denoted as

$$\begin{aligned} &\{t_{ij}(r), i = 1, \ldots, n-1, j = i+1, \ldots, n\} \\ &\quad = \{\tau_l(r), l = 1, \ldots, L\} \end{aligned} \tag{9}$$

Therefore, the values of all the transmissibility functions can be written as

$$\Gamma = \begin{bmatrix} \tau_1(1) & \tau_1(2) & \ldots & \tau_1(R) \\ \tau_2(1) & \tau_2(2) & \ldots & \tau_2(R) \\ \vdots & \vdots & \vdots & \vdots \\ \tau_L(1) & \tau_L(2) & \ldots & \tau_L(R) \end{bmatrix} \tag{10}$$

For the purpose of damage detection, the transmissibility correlation (TC) between healthy ${}^h\tau(r)$ and in-service $\tau(r)$ can be used, which is defined as

$$TC(r) = \frac{\left| \sum_{l=1}^{L} \tau_l(r)^h \tau_l(r) \right|^2}{\left[\sum_{l=1}^{L} \tau_l(r) \tau_l(r) \right] \left[\sum_{l=1}^{L} {}^h\tau_l(r)^h \tau_l(r) \right]} \tag{11}$$

Transmissibility damage indicator (TDI) is the average of transmissibility correlations over all the parameters $\{1, \ldots, R\}$ given

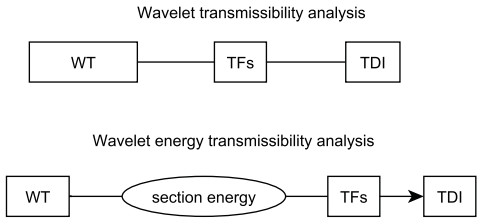

Fig. 3. Wavelet section energy transmissibility procedure.

by [34]

$$TDI = \frac{1}{R}\sum_{r=1}^{R} TC(r) \tag{12}$$

The wavelet TDI measures the similarities between normal condition and monitored condition. If the two conditions are similar, the correlation values at all frequencies are high. Otherwise, the correlation values are low. It is worth mentioning that the range of wavelet TDI values is between 0 and 1. If wavelet TDI value is 1 or near 1, it means the monitored condition is healthy. If it is smaller than 1 or near 0, it indicates damage may happen. In general, the more serious the damage is, the smaller the TDI value will be [21].

III. Wavelet Energy Transmissibility Function (WETF)

In theory, the WTF should be unique under the same system condition. However, in practice, the WTF is determined from measured data and the results can therefore be affected by many issues. First, the WTF can be easily corrupted by the noise due to that a ratio calculation for estimating WTF is used. Further, under different loading conditions, the WTF often varies because of the loading effects. For example, a wind turbine operates under time-varying wind loads, and therefore it has time-varying system responses, say vibrations. As WTF is estimated using measured vibration data in practice, the estimated values may not be consistent under different wind loads. In other words, the effects of time-varying wind loads may not be fully eliminated. These two reasons may prevent the wide applications of WTF. In order to address these issues, in this section, a new concept, called as wavelet energy transmissibility function (WETF), is proposed. Unlike the WTF that is the ratio of the coefficients between wavelet transforms of two different responses [35], [36], the WETF is the ratio of wavelet energy between the wavelet transforms of two different responses. Fig. 3 illustrates the difference between WTF and WETF. Here, the wavelet energy is defined as the root mean squares (RMS) of a group of wavelet coefficients. Suppose the R wavelet coefficients are divided into Z groups and each group has m wavelet coefficients. More specifically, from the original wavelet coefficients in (6), the new wavelet energy (WE) can be obtained as

$$WE_i = \{e_i(z), z = 1, \ldots, Z\} \tag{13}$$

where

$$\begin{bmatrix} e_i(1) = RMS(\nu_i(1), \ldots, \nu_i(m)) \\ e_i(2) = RMS(\nu_i(m+1), \ldots, \nu_i(2m)) \\ \vdots \\ e_i(Z) = RMS(\nu_i((Z-1)m+1), \ldots, \nu_i(Zm)) \end{bmatrix} \tag{14}$$

and

$$\begin{aligned} e_i(z) &= RMS(\nu_i((z-1)m+1), \ldots, \nu_i(zm)) \\ &= \sqrt{\frac{1}{m}\sum\nolimits_{h=(z-1)m+1}^{zm} \nu_i^2(h)} \end{aligned} \tag{15}$$

and $zm = R$. Then the WETF is defined as

$$T_{ij} = [t_{ij}(1), \ldots, t_{ij}(z), \ldots, t_{ij}(Z)] \tag{16}$$

where

$$t_{ij}(z) = \frac{e_i(z)}{e_j(z)} \tag{17}$$

The total number of the WETF is also $L = (n-1)n/2$ and these functions will be represented as:

$$\begin{aligned} &\{t_{ij}(z), i = 1, \ldots, n-1, j = i+1, \ldots, n\} \\ &\quad = \{\tau_l(z), l = 1, \ldots, L\} \end{aligned} \tag{18}$$

In this case, the transmissibility correlation (TC) is changed to

$$TC(z) = \frac{\left|\sum_{l=1}^{L} \tau_l(z)^h \tau_l(z)\right|^2}{\left[\sum_{l=1}^{L} \tau_l(z)\tau_l(z)\right]\left[\sum_{l=1}^{L} {}^h\tau_l(z)^h \tau_l(z)\right]} \tag{19}$$

and TDI is the average of the TCs over all the Z groups given by

$$TDI = \frac{1}{Z}\sum_{z=1}^{Z} TC(z) \tag{20}$$

The ultimate objective of grouping wavelet coefficients is to make WTF less sensitive to variable loadings and noise. In each group, its RMS value is the wavelet energy indicator, which is a robust and stable indicator in condition monitoring, and can help reduce sensitivity to noise and dynamic loadings without sacrificing sensitivity to the changes in system properties. The grouping number m can be tuned to control the tradeoff between sensitivities and robustness.

To make the new concept clear and demonstrate its robustness to noise and time varying loading, a numerical example of a mass-damping-spring system is used to show its relationship with conventional TF and advantages over WTF. For the mass-damping-spring system, the TF is only dependent on the mass, damping and spring parameters. Here, suppose one TF that describes the relationship between two measurements in Laplace transform is given by [37]

$$TF(s) = \frac{36s + 400}{s^2 + 36s + 400} \tag{21}$$

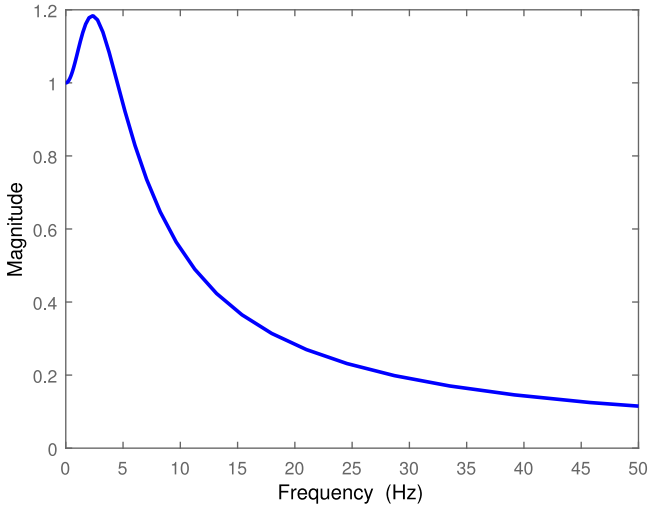

Fig. 4. Transmissibility function of the system used in the numerical example.

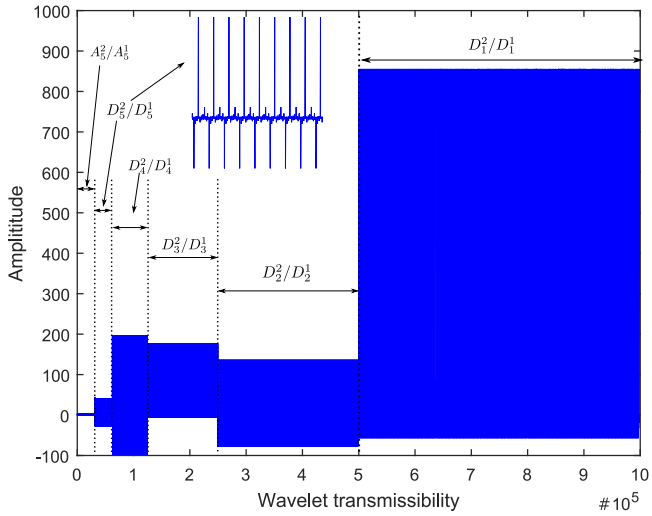

Fig. 5. Wavelet transmissibility function of the system used in the numerical example.

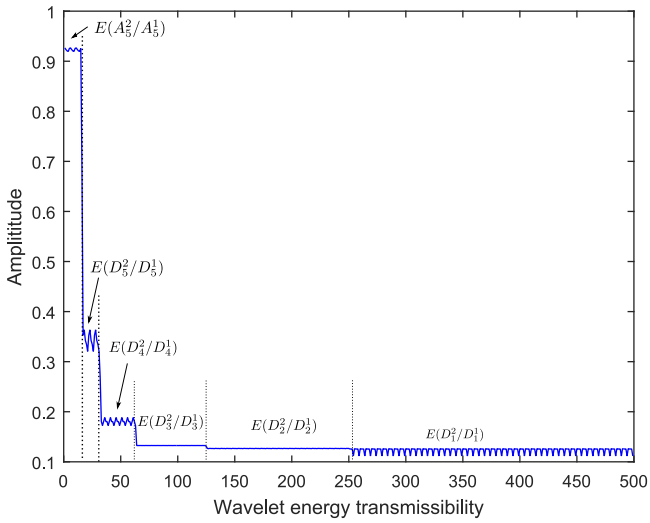

Fig. 6. Wavelet energy transmissibility function of the system used in the numerical example.

where s represents the frequency operator in the Laplace domain. The plot of the frequency response function of the TF is shown in Fig. 4. The magnitude represents the frequency gain of the second measurement over the first measurement. It can be seen that the low frequency gains are around 1 and the high frequency gains fall rapidly. Moreover, suppose the first measurement contains a combination of sine waves over the frequency range from 0 Hz to 50 Hz, with the difference between two consecutive frequencies being 0.2 Hz and the second measurement is the response of the system represented by the TF to the first measurement. To produce the wavelet based TF, the wavelet function and the decomposition level have to be chosen first. Here, the low pass wavelet filter in (5) and high pass filter in (4) are chosen as $h = [\{0, 0, 0, 0, 0.1768, 0.5303, 0.5303, 0.1768, 0, 0, 0, 0\}]$ and $g = [0.0138, 0.0414, -0.0525, -0.2679, 0.0718, 0.9667, -0.9667, -0.0718, 0.2679, 0.0525, -0.0414, -0.0138]$ [38], respectively, and 5 level decomposition is used. Then the WTF and WETF can be calculated using the formula (7) and (16), respectively. Finally, the following comparisons are made.

- *Relationship with TF:* With the 5 level decomposition, the wavelet transform (WT) of each measurement produce 6 groups of wavelet coefficients, namely, $D_1, D_2, D_3, D_4, D_5, A_5$ where D_1 group contains the highest frequency components and A_5 denotes the lowest frequency components and others represent middle frequencies. Further, the wavelet transforms are essentially the filters and therefore the wavelet coefficients are the filtered results of original signal followed by down sampling. In other words, the wavelet coefficients from each group can be regarded as the filtered time-series data. Different from the TF plot that shows the frequency response, both WTF and WETF plots as shown in Figs. 5 and 6 can be classified into six different groups where each contains multiple frequency components. Further, within each group, both of WTF and WETF have repeatable patterns under noise-free conditions and time-invariant loadings. More importantly, WETF has a high ratio in the lowest frequency group represented by $E(A_5^2/A_5^1)$ and lower ratios in other higher frequency groups, which shares the similar trends. This is because the wavelet energy ratio can be interpreted as an average of the mean squared TF frequency gains, which is proved as follows. Suppose $\mathbf{V}_1 = [\nu_1((z-1)m+1), \ldots, \nu_1(zm)]$ and $\mathbf{V}_2 = [\nu_2((z-1)m+1), \ldots, \nu_2(zm)]$ are wavelet transforms of the two measurements and they belong to the defined wavelet energy group $z, z = 1, \ldots, Z$. Since the wavelet coefficients from each level decomposition can be interpreted as the filtered measurement data and the frequency components in a bounded frequency range, $\mathbf{V}_1$ used in the zth group of WETF can be written in a Fourier Series form, i.e.,

$$\nu_1(h) = \sum_{k=lf}^{sf} x_k e^{-i2\pi kh} \tag{22}$$

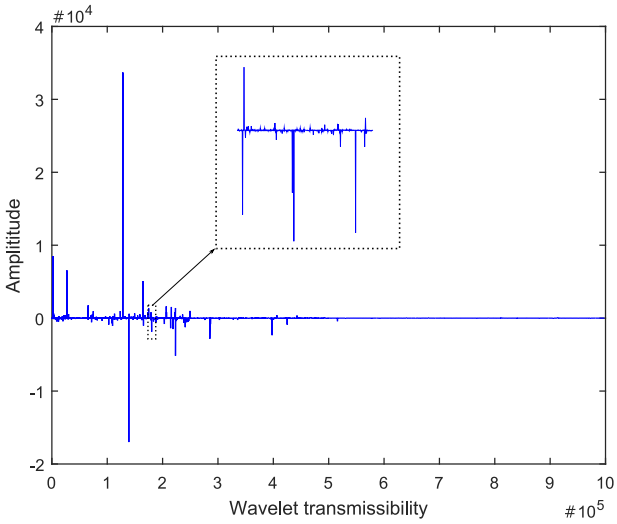

Fig. 7. Wavelet transmissibility function of the system used in the numerical example under noise with 20 SNR.

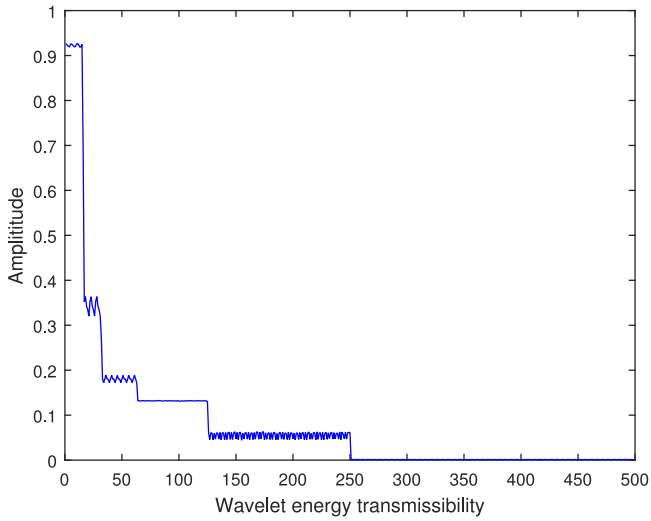

Fig. 8. Wavelet energy transmissibility function of the system used in the numerical example under noise with 20 SNR.

where lf and sf represent the lower bound and upper bound frequency components, respectively, and $h = (z-1)m+1, \ldots, zm$. Since $\nu_2(h)$ can be treated as the output of TF under the input of $\nu_1(h)$, it can be written as

$$\nu_2(h) = \sum_{k=lf}^{sf} c_k x_k e^{-i2\pi kh} \tag{23}$$

where c_k is the gain at the frequency k. By using Parseval's theorem [39], the proposed WETF given by (16) with two measurements can be re-written as

$$\begin{aligned} t(z) = \frac{e_2(z)}{e_1(z)} &= \sqrt{\frac{\sum_{h=(z-1)m+1}^{zm} \nu_2^2(h)}{\sum_{h=(z-1)m+1}^{zm} \nu_1^2(h)}} \\ &= \sqrt{\frac{\sum_{k=lf}^{sf} c_k^2 x_k^2}{\sum_{k=lf}^{sf} x_k^2}} \end{aligned} \tag{24}$$

As $t(z)$ is in a limited frequency range and bounded within $[lf, sf]$, all the gains c_k's are similar and it is reasonable to suppose all the gains can be approximated using a single value $c \in [c_{\min}, c_{\max}]$ where $c_{\min}$ and $c_{\max}$ represent the minimal and maximal gains within the frequency range $[lf, sf]$, respectively. And then the formula (24) can be simplified to

$$\begin{aligned} t(z) = \frac{e_2(z)}{e_1(z)} &= \sqrt{\frac{\sum_{k=lf}^{sf} c_k^2 x_k^2}{\sum_{k=lf}^{sf} x_k^2}} \\ &= \sqrt{\frac{\sum_{k=lf}^{sf} c^2 x_k^2}{\sum_{k=lf}^{sf} x_k^2}} \\ &= c \end{aligned} \tag{25}$$

This indicates that the proposed WETF can be interpreted as a gain of TF within the corresponding frequency range.

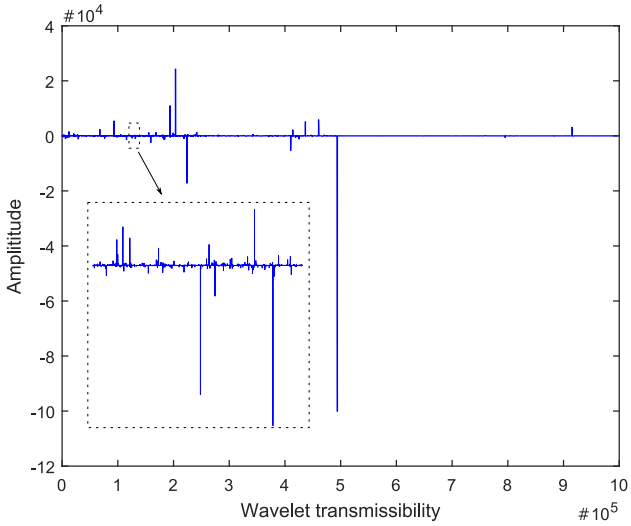

Fig. 9. Wavelet transmissibility function of the system used in the numerical example under varying loading conditions.

In other words, the proposed WETF has a clear physical meaning.

- *Robustness to noise:* The noise with 20 dB signal-to-noise (SNR) ratio is added to the first measurement. The noisy first measurement passes through the TF and this results in the noisy second measurement. Both the WTF and the WETF under the noisy condition are plotted in Figs. 7 and 8, respectively. Compared to noise-free case, the WTF is significantly changed with noise, particularly, repeatable patterns are corrupted. However, the WETF keeps the same trends as those in the noise free case. In other words, the WTF is very sensitive to noise but the WETF is robust to noise.
- *Robustness to time-varying loading:* For a non-stationary process, varying loading has a directly impact on each measurement. Here, suppose the loading only changes once at the half time and therefore two loading conditions are

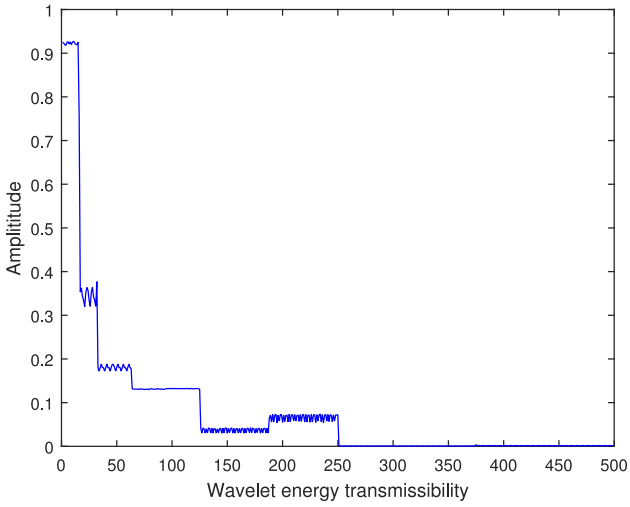

Fig. 10. Wavelet energy transmissibility function of the system used in the numerical example under varying loading conditions.

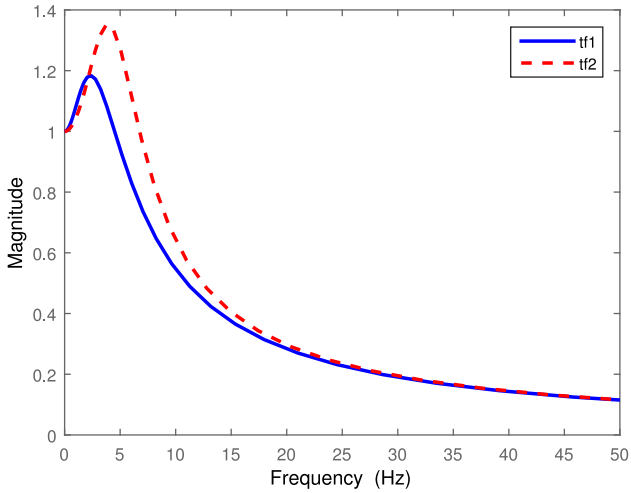

Fig. 11. Two transmissibility functions of the system used in the numerical example.

produced. The first half measurement is under one loading as mentioned above and the latter is under another loading, say a doubled amplitude loading. Results from Figs. 9 and 10 show that varying loading can change the WTF but it has negligible impact on the WETF. Therefore, the WETF is insensitive to varying loading and can be used for non-stationary applications.

- *Sensitivity to system property change:* To test the sensitivity of both WTF and WETF under the condition of system property changes, the experiment is repeated but the system property is changed under the second loading condition. More specifically, the system under two loading conditions has two different TFs that are shown in Fig. 11 where the first one labeled tf1' is given by (21) and the second one labeled by 'tf2' is given by

$$TF(s) = \frac{36s + 900}{s^2 + 36s + 900} \tag{26}$$

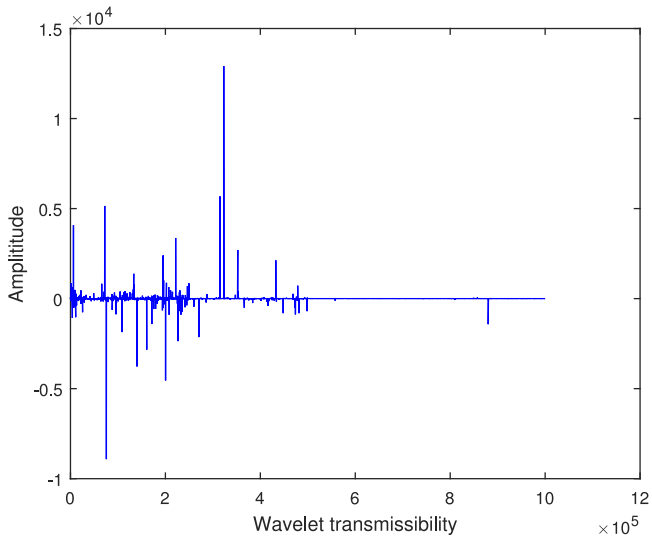

Fig. 12. Wavelet transmissibility function of the system with property change used in the numerical example.

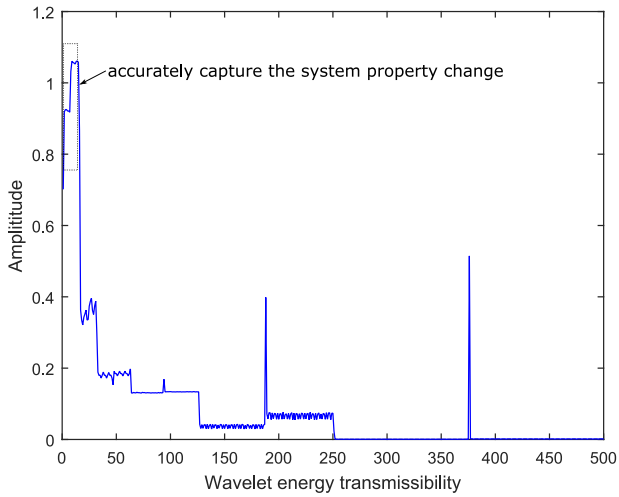

Fig. 13. Wavelet energy transmissibility function of the system with property change used in the numerical example.

The main differences between the two TFs lie in the low frequency gains. It is desirable to capture the system change under different loading conditions. It seems hard to capture the change from the WTF plot shown in Fig. 12. In contrast, it is easy to observe the change from the WETF plot in Fig. 13. WETF correctly indicates the change in the first group with low frequency range that is labeled using a rectangle. Further, as the results of wavelet transforms include time information, it can be seen from Fig. 13 that the change happens in the second half part, which accurately captures the change due to the second loading.

In order to fully understand the WETF, some discussions are summarized as follows:

- *Group number determination:* It is important to know how to determine the total group number Z. Both time-varying

loads and noise have an impact on the group number. More specifically, time-varying or non-stationary loads can be formulated as a combination of a number of time-invariant or stationary loads, say the number is Z_0 under the given sampling time. The basic requirement is that each group represents a full or part of stationary process. Therefore, the chosen group number Z should be bigger than Z_0, e.g. $Z > Z_0$. Further, each group should include sufficient data points where the noise, generally Gaussian noise, can have minimal impact on each group, and then the RMS values would be consistent. On the other hand, the sufficient data in each group requires that group number should not be bigger than an upper bound value, say Z_1. In some cases, Z_0 could be known. For example, wind is time varying but statistical study shows it can be treated a stationary process in a short time of 10 seconds. If the measurement sampling time lasts for 100 seconds, then the total group number is at least 10 in order to make sure that each group represents a stationary process. In some practical cases, Z_0 and Z_1 may be unknown. Therefore, trials and errors can be used to choose the group number Z.

- *WTF is a special case of WETF:* When applying the WETF, if the group number is chosen as its maximal value, i.e., the length of the whole measured data and each group of WETF only has one data, it can be easily to observe that the WETF becomes the WTF. Therefore, the WTF is a special case of the WETF. The WTF has only one single data in each group. This means that it has extremely insufficient data in each group, which results in the high sensitivity to noise. This conclusion is also confirmed using both numerical examples in the current section and the real world case study in the following section.
- *Reference:* In practice, if no healthy condition reference is available, the proposed method can still be used. The current state, even with some defects, can be used as the reference and further deterioration causing more severe defects can be shown in the TDI values.
- *Incipient fault detection and prognosis:* It is very important to detect the incipient faults at an early stage. However, conventional spectral methods often fail to do so under the time-varying loading conditions. The reason for this is that the time-varying loading can lead to varying spectra. The incipient faults with slight changes in the spectra may not be identifiable. The proposed methods can effectively deal with the varying loading impact and therefore it can be used for incipient fault detection. The TDI indicators using the WETF can be used for inferring both damage and severity level detection. Further, the trending of TDI indicators can be used for prognosis to predict the future deterioration rate and even failure time by using some curve analysis methods. It is worth pointing out that, to monitor the incipient faults, data collection should last for a long enough periods, sometimes several years, including both healthy condition and changed conditions. If the proposed method is applied for the continuously monitored data, TDI indicator values will be reduced, often slowly drifted due to the low change in the condition, which can be used to distinguish the incipient damage from the healthy condition because the used TDI indicator is a quantitative evaluation of differences between the reference and in-serve condition.
- *Differences from existing wavelet methods:* First, the unique contribution of this paper is that the novel method does not require or measure loading information but it is able to remove the impact of varying loadings for non-stationary applications. Some popular methods reported in [40] have to measure and use information such as rotation speed in order to deal with time varying conditions. Further, the proposed method is computationally efficient as its main computation is from wavelet decomposition and it does not require a complex training process. A number of existing methods which are a combination of the extracted wavelet features and artificial intelligent algorithms [41], [42], require a large amount training data and a complex training process.
- *Advantages and limitations:* The main advantage is that the new WETF is related to system physical properties and therefore can present the system conditions. Further, the WETF is able to deal with multiple system responses without requiring the system input. Finally, the WETF is computationally efficient as it only needs calculating wavelet decomposition and RMS. The main limitation, as pointed in [27], [43], is that the values of TF depend on location of system input. In this case, multiple WETFs under inputs at different locations can be used to address the varying position problem [34], which will be investigated in future studies.

In this section, the novel concept, WETF, has been introduced and its robustness to dynamic loading and noise has been analyzed in theory and also validated using a numerical example. The application of the novel method to the real-world problem will be investigated in the following section.

IV. Wind Turbine Bearing Condition Monitoring

Worldwide installed wind turbines have been significantly increased over the past decade. To avoid unexpected failure and minimize turbine downtime, different wind turbine condition monitoring systems and methods have been developed [44]–[48]. Although a number of methods have been proposed, most of them are tested in simulation or lab stage and have not been fully tested in operating wind turbines [3]. In this section, the real world wind turbine condition monitoring problem is considered. Two condition monitoring systems were installed on two operating turbines in Greece, which were carried out by an industrial partner in a joint project funded by European Research Council. The vibration data from four acceleration sensors fitted on the main bearing were collected. It is known that one bearing was in good conditions while the other had some damaging conditions over the period of monitoring. It has to mention that for the purpose of condition monitoring, data collection should last for a long enough period including both healthy condition and changed conditions. However, often in practice, no reference data were recorded during the healthy condition. To deal with this issue, an alternative option is to use the healthy data from another system that has the same physical structures but is

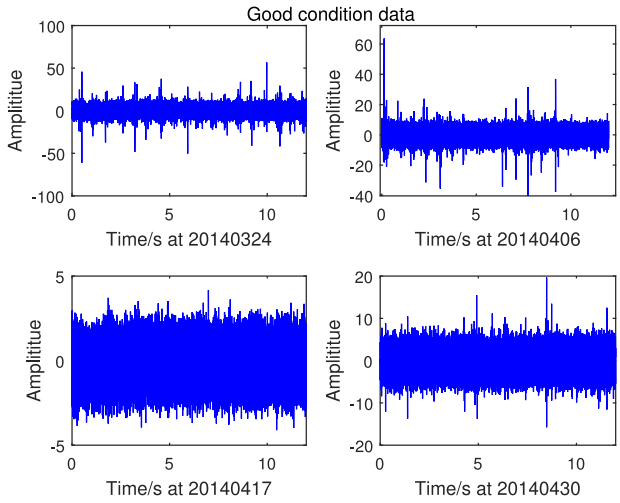

Fig. 14. Time series vibration data of good condition collected at different times.

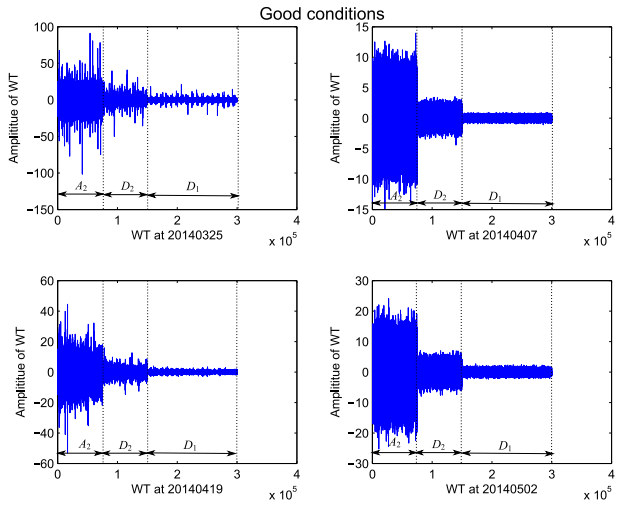

Fig. 16. WT of good condition data collected at different times.

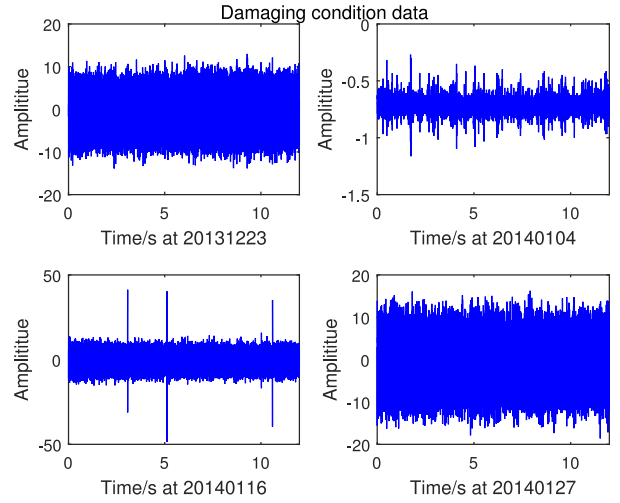

Fig. 15. Time series vibration data of damaging condition collected at different times.

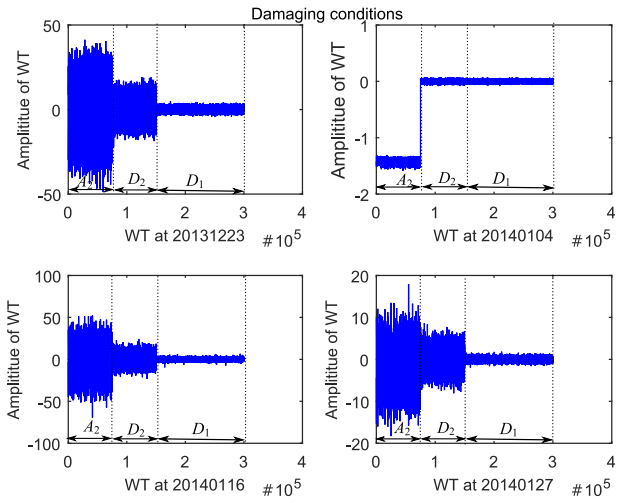

Fig. 17. WT of bad condition data collected at different times.

in good condition. A number of examples in the literature also used another system as a reference [9], [10], [13], [21], [49], [50]. The four acceleration type vibration sensors were fitted at different locations along the main bearing. The employed sampling rate was 25 KHz and each data collection lasted for 12 seconds, producing 300000 data points from each sensor. The data acquisition was carried out hourly and had a duration of about 5 months.

In order to show the time-varying loading effects, four data sets that were collected at different time under good conditions were plotted in Fig. 14, where the data collection date was labeled below each sub-figure, e.g. 20140324 representing 24th March 2014. It can be seen that the four sub-figures have different amplitudes ranges, indicating different loading conditions. Similar phenomena can be observed in Fig. 15 where the data were collected under damaging conditions. Further, the wavelet transforms (WT) of good condition and damaging condition data collected at different times are plotted in Figs. 16 and 17, respectively, where the low pass filter in (5) and high pass filter in (4) are chosen as $h = [0, 0, 0, 0, 0.1768, 0.5303, 0.5303, 0.1768, 0, 0, 0, 0]$ and $g = [0.0138, 0.0414, -0.0525, -0.2679, 0.0718, 0.9667, -0.9667, -0.0718, 0.2679, 0.0525, -0.0414, -0.0138]$ [38], respectively. It can be seen that the wavelet coefficients varies from one sub-figure to another, which also shows the time-varying loading effects. However, the differences between good and bad conditions can not be observed from the wavelet coefficients due to their variations. Finally, the WTFs under good and damaging conditions are shown in Figs. 18 and 19, respectively. It is obviously still hard to tell the differences between good and damaging condition from the WTF due to the variations. Further, the TDI results using WTF are shown in Fig. 20, showing that the good condition has some difference compared with the bad conditions. However, the differences between two conditions are very insignificant. TDI values for the first 2190 sets of good condition data are around 0.3. As the TDI values are in the range of [0,1], we need to choose a threshold to determine its condition. Here, if we choose 0.5 as the threshold value, the

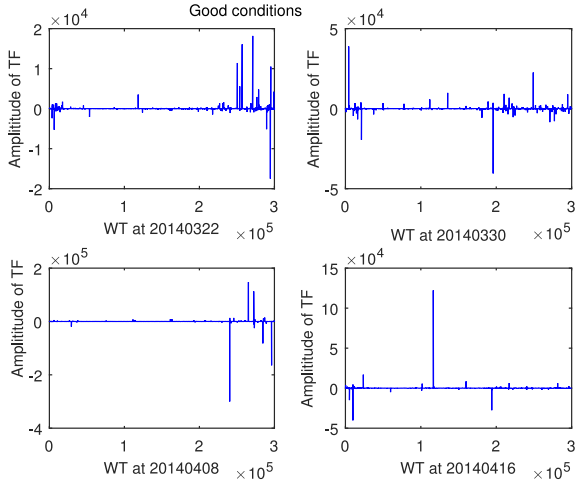

Fig. 18. WTFs of good condition data collected at different times.

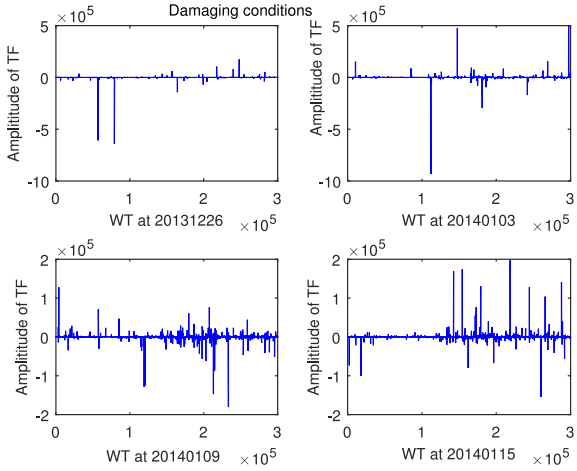

Fig. 19. WTFs of bad condition data collected at different times.

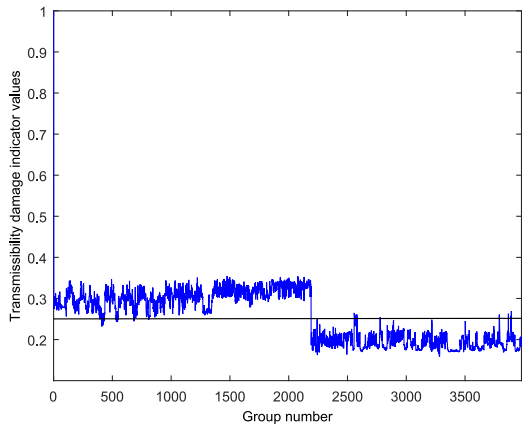

Fig. 20. TDI results using WTF method where data sets (1-2190 datasets: good condition, 2191–3990 datasets: damaging condition).

good condition can be easily misinterpreted as bad conditions due to the small TDI values.

The new WETF is then used to analyze the same data sets. The grouped wavelet coefficients of good and damaging condition data collected at different times is plotted in Figs. 21

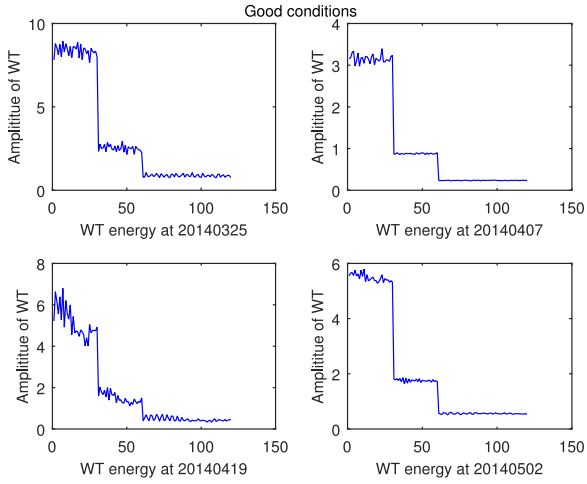

Fig. 21. RMS values of the grouped WT coefficients of good condition data collected at different times.

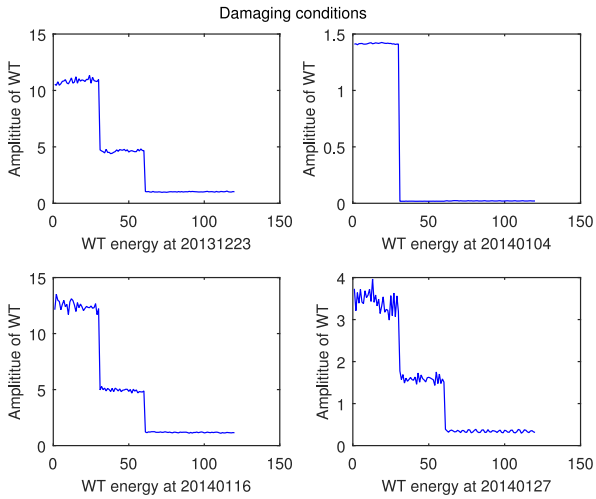

Fig. 22. RMS values of the grouped WT coefficients of bad condition data collected at different times.

and 22, respectively, where the number of wavelet m in each group is chosen as 30 by trial-and-error. It is worth pointing out that the choice of m for other applications should follow the suggestion given in Section III. It can be seen that the amplitudes of grouped wavelet coefficients varies from one sub-figure to another sub-figure, still indicating different loading conditions on different data collection. Further, the WETF under good and damaging conditions are shown in Figs. 23 and 24, indicating the WETFs with good condition under different loading conditions are quite similar. In other words, the WETF is robust to variable loadings and noise. Therefore, it can be used as a good indicator of true system condition. The same conclusion can also be observed from WETFs for the data under damaging condition at different loading conditions. Finally, TDI result using WETF is given in Fig. 25. Compared to the previous results produced by WTF shown in Fig. 20, the new results shown in Fig. 25 can clearly distinguish the differences between good and damaging conditions without a false alarm if choosing 0.5 as the threshold. To make a fair comparison, the

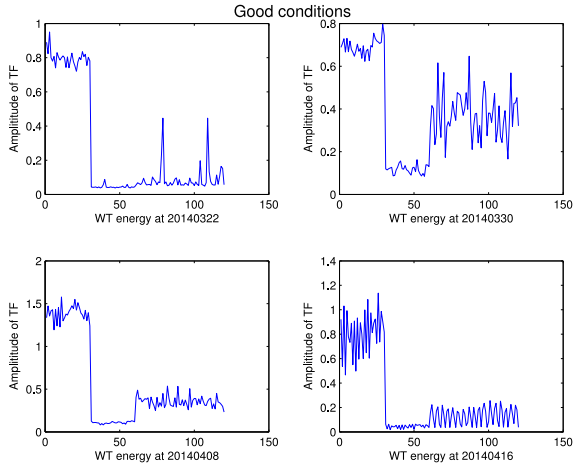

Fig. 23. WETF of good condition data collected at different times.

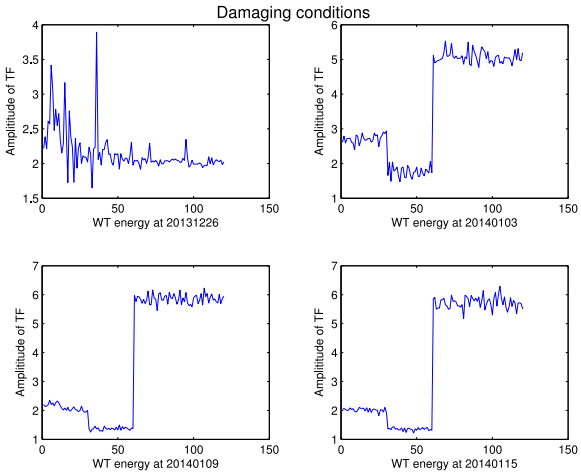

Fig. 24. WETF of bad condition data collected at different times.

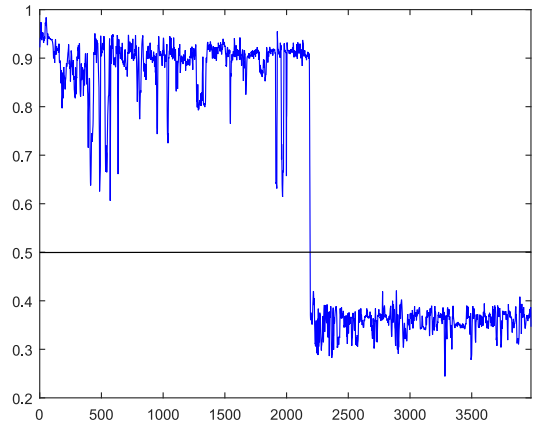

Fig. 25. TDI results using the proposed WETF method (1-2190 datasets: good condition, 2191–3990 datasets: damaging condition).

threshold for the conventional WTF method has been chosen as 0.25 as all its TDI values as shown in Fig. 20 are smaller than those for the proposed method. It is found that there are 85 false alarms for the WTF method. This comparison confirms the advantages of the new WETF approach over the WTF method.

V. Conclusion

In this paper, a new wavelet energy transmissibility analysis method is proposed and has been applied to field data from operating wind turbines for wind turbine bearing condition monitoring. The main advantage of the WETF is that it can be used for condition monitoring under non-stationary operations and the result is insensitive to time-varying loads and noise. The results of the analysis using the field data have shown the new method can produce much better indicator than the direct wavelet transmissibility analysis for evaluating the conditions of wind turbine bearings.

Acknowledgment

The authors would like to thank D2S international (http://www.d2sint.com/) for kindly providing the field data.

References

[1] F. P. G. Márquez, A. M. Tobias, J. M. P. Pérez, and M. Papaelias, "Condition monitoring of wind turbines: Techniques and methods," *Renew. Energy*, vol. 46, pp. 169–178, 2012.

[2] W. Qiao and D. Lu, "A survey on wind turbine condition monitoring and fault diagnosis-part ii: Signals and signal processing methods," *IEEE Trans. Ind. Electron.*, vol. 62, no. 10, pp. 6546–6557, Oct. 2015.

[3] W. X. Yang, P. J. Tavner, C. J. Crabtree, Y. Feng, and Y. Qiu, "Wind turbine condition monitoring: Technical and commercial challenges," *Wind Energy*, vol. 17, no. 5, pp. 673–693, 2014.

[4] B. Lu, Y. Li, X. Wu, and Z. Yang, "A review of recent advances in wind turbine condition monitoring and fault diagnosis," in *Proc. IEEE Power Electron. Mach. Wind Appl.*, 2009, pp. 1–7.

[5] W. Bartelmus and R. Zimroz, "A new feature for monitoring the condition of gearboxes in non-stationary operating conditions," *Mech. Syst. Signal Process.*, vol. 23, no. 5, pp. 1528–1534, 2009.

[6] S. Cruz, "An active-reactive power method for the diagnosis of rotor faults in three-phase induction motors operating under time-varying load conditions," *IEEE Trans. Energy Convers.*, vol. 27, no. 1, pp. 71–84, Mar. 2012.

[7] R. R. Schoen and T. G. Habetler, "Evaluation and implementation of a system to eliminate arbitrary load effects in current-based monitoring of induction machines," *IEEE Trans. Ind. Appl.*, vol. 33, no. 6, pp. 1571–1577, Nov./Dec. 1997.

[8] R. Zimroz, W. Bartelmus, T. Barszcz, and J. Urbanek, "Diagnostics of bearings in presence of strong operating conditions non-stationarity—A procedure of load-dependent features processing with application to wind turbine bearings," *Mech. Syst. Signal Process.*, vol. 46, no. 1, pp. 16–27, 2014.

[9] W. Bartelmus and R. Zimroz, "Vibration condition monitoring of planetary gearbox under varying external load," *Mech. Syst. Signal Process.*, vol. 23, no. 1, pp. 246–257, 2009.

[10] W. Bartelmus, F. Chaari, R. Zimroz, and M. Haddar, "Modelling of gearbox dynamics under time-varying nonstationary load for distributed fault detection and diagnosis," *Eur. J. Mech.-A/Solids*, vol. 29, no. 4, pp. 637–646, 2010.

[11] I. Antoniadou, G. Manson, W. J. Staszewski, T. Barszcz, and K. Worden, "A time–frequency analysis approach for condition monitoring of a wind turbine gearbox under varying load conditions," *Mech. Syst. Signal Process.*, vols. 64–65, pp. 188–216, 2015.

[12] X. Y. Wang, V. Makis, and M. Yang, "A wavelet approach to fault diagnosis of a gearbox under varying load conditions," *J. Sound Vibration*, vol. 329, no. 9, pp. 1570–1585, 2010.

[13] R. B. Randall and J. Antoni, "Rolling element bearing diagnostics—A tutorial," *Mech. Syst. Signal Process.*, vol. 25, no. 2, pp. 485–520, 2011.

[14] E. S. Carbajo, R. S. Carbajo, C. McGoldrick, and B. Basu, "Asdah: An automated structural change detection algorithm based on the Hilbert–Huang transform," *Mech. Syst. Signal Process.*, vol. 47, no. 1, pp. 78–93, 2014.

[15] S. Nagarajaiah and B. Basu, "Output only modal identification and structural damage detection using time frequency & wavelet techniques," *Earthquake Eng. Eng. Vibration*, vol. 8, no. 4, pp. 583–605, 2009.

[16] B. Basu, S. Nagarajaiah, and A. Chakraborty, "Online identification of linear time-varying stiffness of structural systems by wavelet analysis," *Struct. Health Monitoring*, vol. 7, no. 1, pp. 21–36, 2008.
[17] D. Cantero and B. Basu, "Railway infrastructure damage detection using wavelet transformed acceleration response of traversing vehicle," *Struct. Control Health Monitoring*, vol. 22, no. 1, pp. 62–70, 2015.
[18] C. Boller, F. Chang, and Y. Fujino, *Encyclopedia of Structural Health Monitoring*. Hoboken, NJ, USA: Wiley, 2009.
[19] H. Zoubek, S. Villwock, and M. Pacas, "Frequency response analysis for rolling-bearing damage diagnosis," *IEEE Trans. Ind. Electron.*, vol. 55, no. 12, pp. 4270–4276, Dec. 2008.
[20] L. Gelman, "The new frequency response functions for structural health monitoring," *Eng. Struct.*, vol. 32, no. 12, pp. 3994–3999, 2010.
[21] L. Zhang, Z. Q. Lang, and M. Papaelias, "Generalized transmissibility damage indicator with application to wind turbine component condition monitoring," *IEEE Trans. Ind. Electron.*, vol. 63, no. 10, pp. 6347–6359, Oct. 2016.
[22] S. N. Ganeriwala, J. Yang, and M. Richardson, "Using modal analysis for detecting cracks in wind turbine blades," *Sound Vibration*, vol. 45, no. 5, pp. 10–13, 2011.
[23] A. M. R. Ribeiro, J. M. M. Silva, and N. M. M. Maia, "On the generalisation of the transmissibility concept," *Mech. Syst. Signal Process.*, vol. 14, no. 1, pp. 29–35, 2000.
[24] M. J. Schulz, A. S. Naser, P. F. Pai, M. S. Linville, and J. Chung, "Detecting structural damage using transmittance functions," in *Proc. SPIE Int. Soc. Optical Eng.*, 1997, pp. 638–644.
[25] T. J. Johnson and D. E. Adams, "Transmissibility as a differential indicator of structural damage," *J. Vibration Acoust.*, vol. 124, no. 4, pp. 634–641, 2002.
[26] T. J. Johnson, R. L. Brown, D. E. Adams, and M. Schiefer, "Distributed structural health monitoring with a smart sensor array," *Mech. Syst. Signal Process.*, vol. 18, no. 3, pp. 555–572, 2004.
[27] S. Chesné and A. Deraemaeker, "Damage localization using transmissibility functions: A critical review," *Mech. Syst. Signal Process.*, vol. 38, no. 2, pp. 569–584, 2013.
[28] K. Worden, G. Manson, and D. Allman, "Experimental validation of a structural health monitoring methodology: Part i. Novelty detection on a laboratory structure," *J. Sound Vibration*, vol. 259, no. 2, pp. 323–343, 2003.
[29] W. X. Yang, Z. Q. Lang, and W. Y. Tian, "Condition monitoring and damage location of wind turbine blades by frequency response transmissibility analysis," *IEEE Trans. Ind. Electron.*, vol. 62, no. 10, pp. 6558–6564, Oct. 2015.
[30] W. J. Staszewski and D. M. Wallace, "Wavelet-based frequency response function for time-variant systems—An exploratory study," *Mech. Syst. Signal Process.*, vol. 47, no. 1, pp. 35–49, 2014.
[31] K. Dziedziech, W. J. Staszewski, B. Basu, and T. Uhl, "Wavelet-based detection of abrupt changes in natural frequencies of time-variant systems," *Mech. Syst. Signal Process.*, vol. 64, pp. 347–359, 2015.
[32] B. Basu, "Identification of stiffness degradation in structures using wavelet analysis," *Construction Building Mater.*, vol. 19, no. 9, pp. 713–721, 2005.
[33] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.
[34] N. M. M. Maia, R. A. B. Almeida, A. P. V. Urgueira, and R. P. C. Sampaio, "Damage detection and quantification using transmissibility," *Mech. Syst. Signal Process.*, vol. 25, no. 7, pp. 2475–2483, 2011.
[35] L. Zhang, Z. Q. Lang, and W. X. Yang, "Wavelet transmissibility analysis for wind turbine blade condition monitoring," in *Proc. Int. Conf. Condition Monitoring Mach. Failure*, 2015, pp. 1–4.
[36] Z. Q. Yang, "Wavelet transmissibility analysis for wind turbine condition monitoring," M.S. thesis, Univ. Manchester, Manchester, U.K., 2016.
[37] Y. E. Lage, M. M. Neves, N. M. M. Maia, and D. Tcherniak, "Force transmissibility versus displacement transmissibility," *J. Sound Vibration*, vol. 333, no. 22, pp. 5708–5722, 2014.
[38] S. Mallat, *A Wavelet Tour of Signal Processing*. San Francisco, CA, USA: Academic, 1999.
[39] D. G. Manolakis and V. K. Ingle, *Applied Digital Signal Processing: Theory and Practice*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
[40] H. H. Bafroui and A. Ohadi, "Application of wavelet energy and shannon entropy for feature extraction in gearbox fault detection under varying speed conditions," *Neurocomputing*, vol. 133, pp. 437–445, 2014.
[41] A. A. Yusuff, C. Fei, A. A. Jimoh, and J. L. Munda, "Fault location in a series compensated transmission line based on wavelet packet decomposition and support vector regression," *Electr. Power Syst. Res.*, vol. 81, no. 7, pp. 1258–1265, 2011.
[42] M. A. Jafarizadeh, R. Hassannejad, M. M. Ettefagh, and S. Chitsaz, "Asynchronous input gear damage diagnosis using time averaging and wavelet filtering," *Mech. Syst. Signal Process.*, vol. 22, no. 1, pp. 172–201, 2008.
[43] D. Christof *et al.*, "Structural health monitoring in changing operational conditions using transmissibility measurements," *Shock Vibration*, vol. 17, nos. 4/5, pp. 651–675, 2010.
[44] L. Stankovic, V.and Stankovic, S. Wang, and S. Cheng, "Distributed compression for condition monitoring of wind farms," *IEEE Trans. Sustain. Energy*, vol. 4, no. 1, pp. 174–181, Jan. 2013.
[45] P. Guo, D. Infield, and X. Yang, "Wind turbine generator condition-monitoring using temperature trend analysis," *IEEE Trans. Sustain. Energy*, vol. 3, no. 1, pp. 124–133, Jan. 2012.
[46] A. Kusiak and A. Verma, "A data-driven approach for monitoring blade pitch faults in wind turbines," *IEEE Trans. Sustain. Energy*, vol. 2, no. 1, pp. 87–96, Jan. 2011.
[47] F. Cheng, L. Qu, and W. Qiao, "Fault prognosis and remaining useful life prediction of wind turbine gearboxes using current signal analysis," *IEEE Trans. Sustain. Energy*, vol. 9, no. 1, pp. 157–167, Jan. 2018.
[48] S. Shokrzadeh, M. J. Jozani, and E. Bibeau, "Wind turbine power curve modeling using advanced parametric and nonparametric methods," *IEEE Trans. Sustain. Energy*, vol. 5, no. 4, pp. 1262–1269, Oct. 2014.
[49] W. Bartelmus and R. Zimroz, "A new feature for monitoring the condition of gearboxes in non-stationary operating conditions," *Mech. Syst. Signal Process.*, vol. 23, no. 5, pp. 1528–1534, 2009.
[50] G. K. Chaturvedi and D. W. Thomas, "Bearing fault detection using adaptive noise cancelling," *J. Mech. Des.*, vol. 104, no. 2, pp. 280–289, 1982.

Long Zhang (M'13) received the B.Eng. and M.Eng. degrees in electrical engineering and automation from Harbin Institute of Technology, Harbin, China, in 2008 and 2010, respectively, and the Ph.D. degree in electronics, electrical engineering and computer science from Queen's University, Belfast, U.K., in 2013. He was a Research Associate with the Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, U.K., from 2014 to 2015. He joined the School of Electrical and Electronic Engineering, University of Manchester, Manchester, U.K., as a Lecturer in 2015. His research mainly focuses on machine and statistical learning, neural networks, system identification, frequency analysis, intelligent control, and their applications to complex systems modeling, analysis, prediction, control, and faut diagnosis. One of his particular interests lies in the wind turbine condition monitoring and fault diagnosis.

Zi-Qiang Lang received the B.S. and M.Sc. degrees in automatic control in China and the Ph.D. degree in systems and control engineering from the University of Sheffield, Sheffield, U.K. He is currently the Chair Professor of complex systems analysis and design with the Department of Automatic Control and Systems Engineering, University of Sheffield. His main expertise relates to the theories and methods for complex systems modeling, analysis, design, signal processing, and the application of these to resolving various science and engineering research problems, including smart structures and systems, civil and mechanical structure vibration control, structural health monitoring, and condition monitoring and fault diagnosis for wind turbine components and systems.