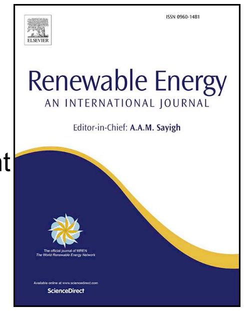
Article:

Zhao, X. and Lang, Z. (2019) Baseline model based structural health monitoring method under varying environment. Renewable Energy, 138. pp. 1166-1175. ISSN 0960-1481

https://doi.org/10.1016/j.renene.2019.02.007

Accepted Manuscript



Baseline model based structural health monitoring method under varying environment

Xueyan Zhao, Ziqiang Lang

PII: S0960-1481(19)30151-X

DOI: https://doi.org/10.1016/j.renene.2019.02.007

Reference: RENE 11139

To appear in: *Renewable Energy*

Received Date: 7 August 2018

Revised Date: 1 January 2019

Accepted Date: 1 February 2019

Baseline model based structural health monitoring method under varying environment

Xueyan Zhao[1], Ziqiang Lang[2]

1. College of Engineering, China Agricultural University, Beijing, 100083, China,

Email: xyzhao@cau.edu.cn

2. Department of Automatic Control and Systems Engineering, University of Sheffield, S1 3JD, UK.

Email: z.lang@sheffield.ac.uk

* Corresponding author. Tel.: +86 (0)10 6273 6665, E-mail: xyzhao@cau.edu.cn.

Abstract: Environment has significant impacts on the structure performance and will change features of sensor measurements on the monitored structure. The effect of varying environment needs to be considered and eliminated while conducting structural health monitoring. In order to achieve this purpose, a baseline model based structural health monitoring method is proposed in this paper. The relationship between signal features and varying environment, known as a baseline model, is first established. Then, a tolerance range of the signal feature is evaluated via a data based statistical analysis. Furthermore, the health indicator, which is defined as the proportion of signal features within the tolerance range, is used to judge whether the structural system is in normal working condition or not so as to implement the structural health monitoring. Finally, experimental data analysis for an operating wind turbine is conducted and the results demonstrate the performance of the proposed new technique.

Keywords: Wind turbine; varying environment; B-spline model; structural health monitoring

1 Introduction

Most structural systems are subject to suffering damage due to inappropriate operation, hostile working conditions or fatigue damage after long time service. Minor damage will change the performance and reliability of the structural system; while serious damage will lead to system malfunction and even cause casualties. Therefore, structural health monitoring (SHM) has been widely employed to monitor structural health status and indicate the possibility of damage in the structural system so that proper maintenance can be scheduled in time to reduce the unexpected loss caused by downtime[1, 2].

Extensive methods have been developed to implement structural health monitoring and fault diagnosis. Model-based and signal-based structural health monitoring methods and their applications were comprehensively reviewed in [3], and knowledge-based and hybrid/active methods were surveyed in [4]. Ma et al. studied different types of damage in rotor systems including rub-impact [5], misalignment [6] and pedestal looseness [7], and experimental results verified the possibility of a finite element method in health monitoring. Especially, many researches focused on the performance of concrete damage-sensitive features. For example, Mohanty et al. [8] investigated vibration of a multistage gearbox with various defects, i.e. one or two teeth broken, and concluded that the input shaft frequency was able to indicate the existence of defects. Williams et al. [9] studied the root mean square (RMS) levels of measurements from an acoustic emission (AE) sensor on inner race of ball bearing, and concluded that the RMS levels of AE sensor measurements exhibited a monotonous increase after the occurrence of damage.

However, the changes revealed by damage-sensitive features which are always considered as SHM features are affected not only by damage in the inspected structural system but also by the working environment [10]. The varying environment has significant impacts on the system dynamic behaviours as discussed by Sohn in [11]. Moreover, Sohn et al. [12] studied the vibration of a theme park ride by combining time series analysis with statistical pattern recognition technique and concluded that the feature variation caused by mass loading was more obvious than that caused by delamination damage. Ha et al. [13] researched the effects of temperature and humidity on pre-stressed concrete girders and found that when the temperature and humidity increased, the frequencies and damping ratios decreased in proportion. The stability of a rotor system with rub-impact damage under different rotating speeds was investigated by Han et al. in [14], and the results revealed that when rotating speed increased, the system exhibited firstly stable, then period-doubling bifurcations and finally reached the stable periodic motion again. As for the gearbox, Loutas et al. [15] researched how the features of the vibration and AE signals in the frequency domain changed when the gearbox kept working until several teeth were cut and considerable damage happened on the shaft. It was concluded that the oil temperature had an effect on the recordings.

Therefore, many researchers have paid attention to the influence of varying environment on system behaviors, and then, try to investigate the effect of non-damage factors so as to enhance the reliability of structural health monitoring methods[16-18]. One type of methods for removing the effects of varying environment is to model the relationship of damage-sensitive features and varying environment. Makis and Yang [19] found that a model developed under the constant load assumption could not recognize whether the vibration feature changes of gearbox were caused by the load variation or by a failure occurrence. To settle this problem, an ARX model was proposed which considered load as additional information. Worden et al. [20] revised the conventional outlier analysis method by replacing the traditional mean vector of damage-sensitive features with features at the same temperature predicted from a polynomial regression model in temperature and the mean vector of damage-sensitive features at this temperature. Zhao and Lang established the relationship between the varying environments and SHM features using a polynomial model [21] and a B-spline model [18] respectively, and then proposed a novel health indicator after removing the environmental effect to indicate health condition of the monitored system. Experimental study on wind turbine components proved the effectiveness of the health indicator. Another type of methods removing the effects of varying environment is to extract signal features which are insensitive to environmental variation but still damage-sensitive. Cross and Worden combined linearly several damage-sensitive features to produce a new feature which was independent of environmental variation but was sensitive to damage in [22], and further tried to extract signal features which were insensitive to environmental variation but still damage-sensitive by co-integration technique, outlier analysis and minor principal components techniques in [23].

Most above researches except [18] and [21] are based on the assumption that the change of SHM features can be generally expressed by the environmental variation within the whole range. But the features of measurements are likely to be influenced obviously by the local environment parameters [18]. Therefore, this paper present a novel and efficient structural health monitoring method by taking environmental variation which is at a similar damage sensitivity level as a group. There are two novelties and contributions in this paper. The first one is that an improved B-spline model is developed to build baseline model between SHM features and environment parameters. This model can deal with local effect very well and fit data smoothly with low degree and high efficiency. The other one is that the structural health monitoring is conducted not in the whole range of environment parameters but in different bins which cover the value of environment parameters at similar damage sensitivity levels, this is benefit to improve the reliability of the structural health monitoring results.

The layout of the paper is as follows. After this introduction, the baseline model based SHM method under varying environment is proposed and demonstrated systematically in Section 2. The effectiveness of the new method is verified by experimental case study in Section 3 and simulation case study in Section 4 respectively. Finally, the conclusions are presented in Section 5.

2 Methodology

Traditionally, structural health monitoring is achieved by monitoring structural signal features and identifying any deviation of these features from a healthy one, an obvious deviation is indicative of a developing damage. The signal feature of the monitored structure can be named as in-service feature, while the signal feature of the healthy structure can be named as health feature. They are extracted respectively from sensor measurements of the monitored structure and the health structure by using a range of data analysis methods [11] , such as time domain analysis, frequency domain analysis or time-frequency domain analysis [9, 24, 25]. However, fluctuating environment has significant impacts on the structure performance, and can also cause the change of signal features which will lead to incorrect results of SHM. In order to remove the effect of fluctuating environment on the results of traditional structural health monitoring, a baseline model is proposed to represent the relationship between healthy SHM features and corresponding environment parameters. Then tolerance ranges of the in-service SHM features under certain environment conditions are obtained by statistical analysis. Finally, in-service structural system health condition can be determined by identifying occurrences of in-service SHM features within tolerance range. Baseline model, tolerance range and health indicator are achieved as follows.

2.1 B-spline based baseline model

The most important part of SHM considering varying environment is the baseline model between healthy SHM features and corresponding environment parameters[26]. The purpose of building a baseline model is to map the system environment parameters to the signal features extracted from the sensor measurements so that the effects of varying environments can be removed when conducting SHM. Baseline model can be expressed as:

$$y = f(x_1, x_2, x_3, \dots, x_M) \tag{1}$$

where $x_1, x_2, x_3, \dots, x_M$ are the environment parameters, M is the number of the environment parameters, and y is the SHM feature. Many methods can be employed to build the baseline model, such as polynomial model [21, 26], ARX model [19] and auto-associative neural network [27]. In this paper, a revised B-spline model is used to determine the baseline model.

Conventional B-spline model can be expressed as [28]

$$y = f(x_1, x_2, x_3, \dots, x_M) = \sum_{i_1=0}^{M_1} \dots \sum_{i_M=0}^{M_M} \alpha_{i_1,i_2,\dots i_M} N_{i_1,p}(x_1) \dots N_{i_M,p}(x_M) \tag{2}$$

where $N_{i_1,p}(x_1)$,…,$N_{i_M,p}(x_M)$ are the i_1^{th},…, i_M^{th} B-spline basis functions of degree p with respect to variables x_1,…, x_M, respectively; and $N_{i_1,p}(x_1)$,…, $N_{i_M,p}(x_M)$ can be expressed by $N_{i_m,p}(x_m)$, $m = 1,2,\dots,M$; $\alpha_{i_1,i_2,\dots i_M}$ is control coefficient of the term $N_{i_1,p}(x_1)\dots N_{i_M,p}(x_M)$; M_m is the number of B-spline basis function of $N_{i_m,p}(x_m)$, where $m = 1,2,\dots,M$. Given a knot vector $\boldsymbol{x}_m = \{x_{m,0}, x_{m,1}, x_{m,2}, \dots, x_{m,K}\}$

and degree p, B-spline basis function $N_{i_m,p}(x_m)$ is usually defined by Cox-de Boor recursion formula as follows:

$$N_{i_m,0}(x_m) = \begin{cases} 1 & \text{if } x_{m,i_m} \le x_m < x_{m,i_m+1} \\ 0 & \text{otherwise} \end{cases} \tag{3.1}$$

$$N_{i_m,p}(x_m) = \frac{x_m - x_{m,i_m}}{x_{m,i_m+p} - x_{m,i_m}} N_{i_m,p-1}(x_m) + \frac{x_{m,i_m+p+1} - x_m}{x_{m,i_m+p+1} - x_{m,i_m+1}} N_{i_m+1,p-1}(x_m) \tag{3.2}$$

It can be deduced from Eqs.(2) and (3) that the basis function $N_{i_m,p}(x_m)$ is non-zero on only $p+1$ knot spans, namely, $[x_{m,i_m}, x_{m,i_m+1})$, $[x_{m,i_m+1}, x_{m,i_m+2})$, …, $[x_{m,i_m+p}, x_{m,i_m+p+1})$, and on any knot span $[x_{m,i_m}, x_{m,i_m+1})$, at most $p+1$ basis functions with degree p are non-zero, namely, $N_{i_m-p,p}(x_m)$, $N_{i_m-p+1,p}(x_m)$, …, $N_{i_m,p}(x_m)$. Thus, changing the control coefficient $\alpha_{i_1,i_2,\dots i_M}$ or the position of knot $x_{m,i}$ only affects the curve shape of B spline model on local span; this is so-called local effect or local modification property. In addition, B-spline curve expressed by Eq.(3) is a piecewise and derivative curve with each component a curve of degree p, this property allows B-spline model to fit complex shapes smoothly with lower degree than ARX model and with higher efficiency than neural network. The B-spline model expressed by Eqs.(2) and (3) has excellent capabilities in smooth data fitting and local effect, and can be used to fit the data with lower degree but higher efficiency, so it is employed in this paper to determine the baseline model.

In order to explain the ability of the B-spline model in fitting the data, one example is provided in the following. Fig.1 shows vibration levels of a rotor system under different rotating speeds, where the horizontal coordinate is the rotating speed of the rotor system with the unit of Hz, and the vertical coordinate is the vibration amplitude of the rotor system with the unit of mm (Detailed information about the rotor system can be found in Case 1 in [29]). The B-spline model is applied to fit data shown in Fig.1. In this case, only rotating speed is treated as an independent variable, so $M = 1$. When the degree of B-spline basis function is set as $p = 3$, the number of knots is set as 15, namely, $K = 14$, and the knot vector is set as

$$\boldsymbol{x}_1 = \{x_{1,0}, \dots, x_{1,K}\} = \{60, 61, 63.3, 66.6, 69.9, 73.2, 76.5, 79.8, 83.1, 86.4, 89.7, 93, 96.3, 99.6, 100\}$$

Then, B-spline basis functions $N_{i_1,p}(x_1)$ can be determined according to Eqs.(3.1) and (3.2), and some of them are shown in Fig. 2. Corresponding coefficients are estimated by least square, and the results are listed in Table 1. The fitting error when $K = 14$ is shown in Fig. 3 by a blue solid line. The maximum, mean and standard deviation are 0.1636, 0.0032 and 0.0582 respectively, indicating that the fitting error is small and ignorable. Therefore, data in Fig.1 can be represented by B-spline model Eqs.(2)-(3) with B-spline basis functions in Fig.2 and corresponding coefficients in Table 1.

Table 1 Coefficients for B-spline model

Terms	Coefficients	Terms	Coefficients	Terms	Coefficients	Terms	Coefficients
$N_{1,3}(x)$	0.3625	$N_{5,3}(x)$	0.9096	$N_{9,3}(x)$	1.7059	$N_{13,3}(x)$	4.0308
$N_{2,3}(x)$	0.2403	$N_{6,3}(x)$	0.8023	$N_{10,3}(x)$	2.1150	$N_{14,3}(x)$	4.7122
$N_{3,3}(x)$	0.6693	$N_{7,3}(x)$	1.1435	$N_{11,3}(x)$	2.6903	$N_{15,3}(x)$	4.9271
$N_{4,3}(x)$	0.3400	$N_{8,3}(x)$	1.4623	$N_{12,3}(x)$	3.1659	$N_{16,3}(x)$	4.9661

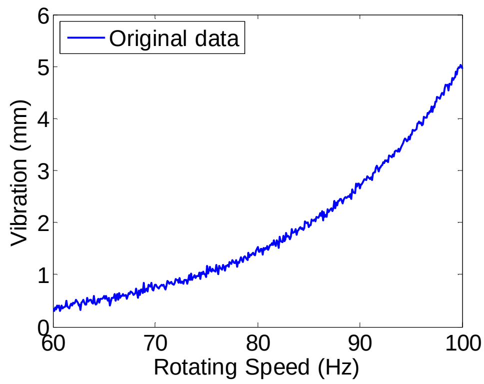


Fig.1 Original data

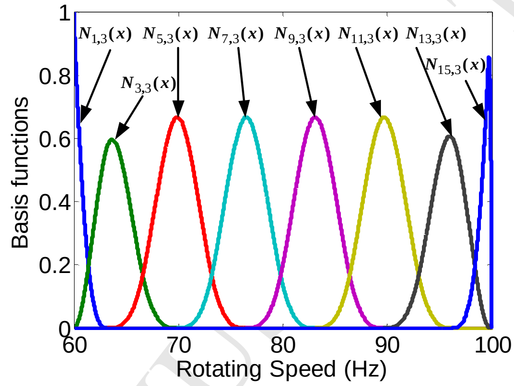


Fig. 2 B-spline basis functions

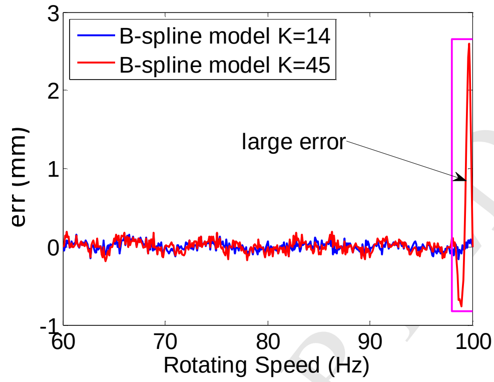


Fig. 3 Fitting error by B-spline model

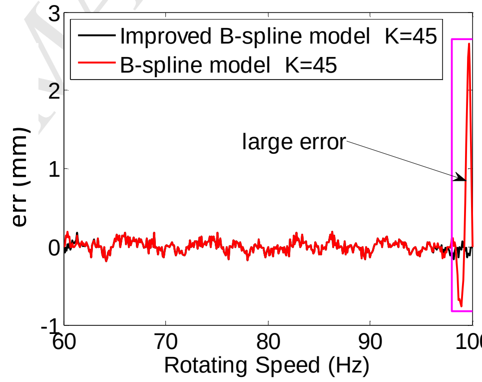


Fig. 4 Fitting error by improved B-spline model

However, when the number of knots K increases, the performance of B-spline model in fitting data becomes unstable, e.g., when the number of knots increases to $K = 45$, fitting error by B-spline model at the end data is much larger than that when $K = 14$ as shown in Fig.3; the maximum, mean and standard deviation are 2.5892, 0.0284 and 0.2725 respectively. This is because corrosion of data at the end tends to deteriorate when the number of knots and B-spline basis functions become larger. Besides, the increase in the number of knots and B-spline basis functions will also lead to more complicated and tedious computations, and computational errors are accumulated when fitting a B-spline model. In order to solve this problem, conventional B-spline model can be improved by reordering all remaining B-spline basis functions and/or ignoring insignificant B-spline basis functions and their multiplications by using recursive forward-regression orthogonal estimator (RFROE)

[30]. The terms which contribute prominently to the model can be selected as follows.

Step (a): All terms $N_{i_1,p}(x_1)\dots N_{i_M,p}(x_M), i_1 = 0,1,2,\dots,M_1,\dots,i_M = 0,1,2,\dots,M_M$ are considered as possible candidates for the most important term $w_1(t)$. For $i_1 = 0,1,2,\dots,M_1,\dots,i_M = 0,1,2,\dots,M_M$, set $w_1^{(i_1\dots i_M)}(t) = N_{i_1,p}(x_1)\dots N_{i_M,p}(x_M)$, then calculate

$$\hat{g}_1^{(i_1\dots i_M)} = \frac{\sum_{t=1}^{N} w_1^{(i_1\dots i_M)}(t)y(t)}{\sum_{t=1}^{N}\left(w_1^{(i_1\dots i_M)}(t)\right)^2} \tag{4}$$

and

$$[err]_1^{(i_1\dots i_M)} = \frac{\left(\hat{g}_1^{(i_1\dots i_M)}\right)^2 \sum_{t=1}^{N}\left(w_1^{(i_1\dots i_M)}(t)\right)^2}{\sum_{t=1}^{N} y^2(t)} \tag{5}$$

Step (b): Find the maximum of $[err]_1^{(i_1\dots i_M)}$, e.g., $[err]_1^{(M_1\dots M_M)} = \max\left\{[err]_1^{(i_1\dots i_M)}, i_1 = 0,1,2,\dots,M_1,\dots,i_M = 0,1,2,\dots,M_M\right\}$. Then the first term is selected with $[err]_1 = [err]_1^{(M_1\dots M_M)}$, and $w_1(t) = w_1^{(M_1\dots M_M)}(t) = N_{M_1,p}(x_1)\dots N_{M_M,p}(x_M)$.

Step (c): All the remaining terms are considered as possible candidates for $w_2(t)$. Set $w_2^{(i_1\dots i_M)}(t) = N_{i_1,p}(x_1)\dots N_{i_M,p}(x_M) - \alpha_{12}^{(i_1\dots i_M)} w_1(t)$, calculate $\hat{g}_2^{(l)}$ and $[err]_2^{(l)}$ by using Eqs.(4) and (5) , respectively, where

$$\alpha_{12}^{(i_1\dots i_M)} = \frac{\sum_{t=1}^{N} w_1(t) N_{i_1,p}(x_1)\dots N_{i_M,p}(x_M)}{\sum_{t=1}^{N} w_1^2(t)} \tag{6}$$

Step (d): Find the maximum of $[err]_2^{(i_1\dots i_M)}$, and then corresponding term $N_{i_1,p}(x_1)\dots N_{i_M,p}(x_M)$ is selected.

Step (e): Then Step (c) and (d) are iterative, and the procedure is terminated at the D_s th step when

$$1 - \sum_{i=1}^{D_s}[err]_i < a\ desired\ tolerance, D_s < D \tag{7}$$

or when $D_s = D$, where D the number of the maximum iterative steps.

The value of the desired tolerance can be determined by using APRESS criteria in [31].

Step (f): Identify coefficients of selected terms, which contribute significantly to the model, by using the least square method.

The fitting error by using improved B-spline model method is shown in Fig.4. The maximum, mean and standard deviation of the fitting error are 0.1885, 0.0009 and 0.0708 respectively, indicating that the value of the fitting error by using improved B-spline model is obviously smaller than that by using conventional B-spline model.

The improved B-spline based model algorithm can be summarized as the flowchart in Fig. 5.

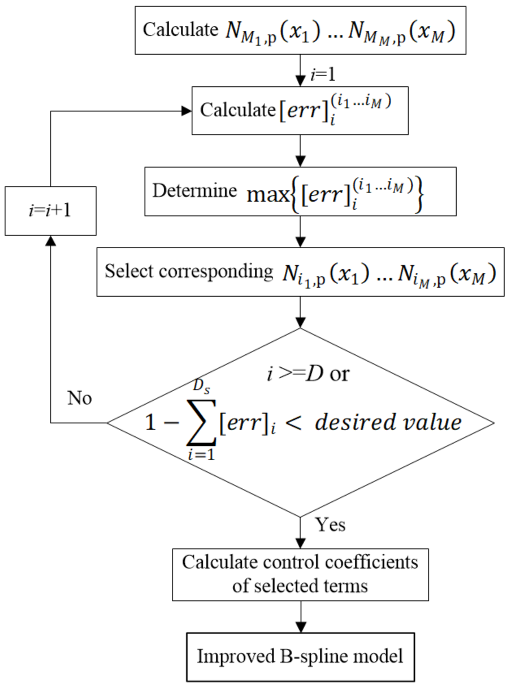


Fig.5 Flowchart of the improved B-spline based model algorithm

2.2 Tolerance range

Denote the deviation between the in-service feature and predicted feature by improved B-spline based baseline model as:

$$e = y' - y \tag{8}$$

where y' is the feature of a sensor measurement from the in-service structural system, y is the feature predicted by the baseline model in Eqs. (2) and (3), e is the deviation between $y'$and y. This deviation is generally determined by many factors, including modelling error, noise, and the effects of less significant environmental changes which cannot be covered by the baseline model. In principle, effects of these factors can be neglected when the structural system is in healthy working conditions, if the baseline model is acceptable in representing the changes of sensor signal features in these conditions. However, damage in the structural system can make a significant change in the deviation, and this phenomenon can be exploited for the structural system health monitoring purpose.

Under the assumption that the deviation e follows a normal distribution when the structural system is working normally, that is, $e \sim N(\mu, \sigma^2)$, where μ and σ are the mean and standard deviation of e, respectively, $[\mu - 3\sigma, \mu + 3\sigma]$ can cover 99.73% of the e values when the structural system is working in healthy conditions. Therefore, the tolerance range of in-service feature y' can be expressed as:

$$y' = y + e \in [y + \mu - 3\sigma, y + \mu - 3\sigma] \tag{9}$$

If y' is within this range, the monitored structural system is working under healthy condition, or else, the monitored structural system is subject to damage in a large degree.

2.3 Health indicator

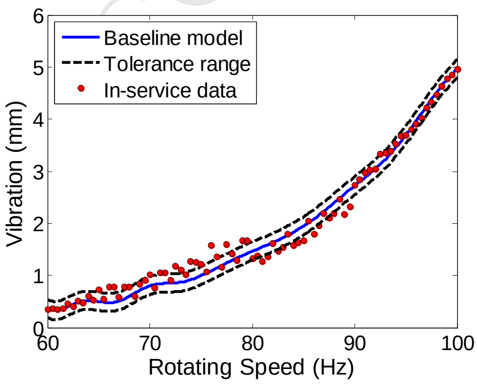
According to the definition of tolerance range above, if a monitored structure is operating in a healthy condition, most in-service y' should fall into the tolerance range. If there is a change or damage, only a small number of values of y' are within the corresponding tolerance range. This phenomenon can be represented quantitatively by the concept of health indicator defined as follows:

$$P = N_{in}/N_{all} \tag{10}$$

where N_{in} is the number of the values of y' where $y' \in [y + \mu - 3\sigma, y + \mu - 3\sigma]$, and N_{all} is the total number of y'.

For example, for data shown in Fig.1, baseline model can be established by using RFROE method in Section 2.1, the obtained improved B-spline model curve is shown as a solid blue line in Fig.6; tolerance range of y' can be calculated by Eq.(9) and shown as a dashed black line in Fig.6; in-service y' is shown as red points in Fig.6. After statistical analysis, total number of y' is 81, 50 of which are within the tolerance range, therefore, health indicator is calculated by Eq.(10) as:

$$P = N_{in}/N_{all} = 50/81 = 0.6173$$



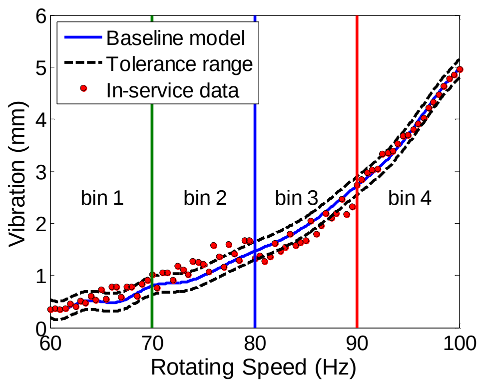


Fig. 6 Tolerance range and in-service data

Fig. 7 Bins

2.4 Health indicator in each bin

The deviation e is likely to vary with the environmental conditions, that is, the value is large in some conditions but small in other conditions. In addition, in practice, signal features of damaged structural systems change slightly in some environmental conditions but change significantly in other environmental conditions. Motivated by these phenomena, the whole environmental conditions are divided into several bins according to the value of environment parameters, so that the deviations which have a similar level can be calculated and their tolerance range can be determined in each bin. The bins can be defined as:

$$B_{n_1 n_2,\dots,n_M} = \{x_1, x_2, \dots, x_M\}, x_i \in \left[x_{i,n_i}, x_{i,n_i+1}\right] \tag{11}$$

where $B_{n_1 n_2,\dots,n_M}$ is the bin when $x_i \in [x_{i,n_i}, x_{i,n_i+1}]$, $i = 1,2,\dots M$; x_{i,n_i} and x_{i,n_i+1} are two edges of n_i th segments for variable x_i; $n_i = 1,2,\dots,M_i$; M_i is the total number of the segments for i th variable x_i. In order to describe bins more precisely, the bins are renumbered by the single subscript.

Tolerance range of in-service feature y' and health indicator can be calculated in each bin separately. For example, for the case shown in Fig.6, when the whole value of x is divided into four bins according to the rotating speeds which cover the range of $x \in [60,70)$, $x \in [70,80)$, $x \in [80,90)$, $x \in [90,100]$, and denoted by Bin 1, Bin 2, Bin 3, Bin 4, respectively as shown in Fig.7. Tolerance range of in-service feature y' in each bin is calculated separately and also shown in Fig.7. N_{in}, N_{all}, and P are calculated in each bin, and the results are shown in Table 2.

Table 2 Calculation of health indicator in each bin

Bin index	N_{in}	N_{all}	P	Bin index	N_{in}	N_{all}	P
Bin 1	16	20	0.80	Bin 3	7	20	0.35
Bin 2	7	20	0.35	Bin 4	20	21	0.9524

2.5 Baseline model based SHM method and remarks

From the above concepts of B-spline based baseline model, bins, tolerance range and health indicator, a new baseline model based SHM method can be proposed. The detailed procedure can be described as follows and summarized as the flowchart in Fig.8.

Step 1: Baseline model establishment: Establish B-spline based baseline model by using RFROE method shown in Eqs.(2)- (7) according to data measured on the healthy structural system;

Step 2: **Bin definition:** Define bins using Eq.(11) according to the value of environment parameters;

Step 3: **Tolerance range calculation:** Calculate tolerance range of SHM feature in each bin using Eq.(9) according to data measured on the healthy structural system;

Step 4: **Health indicator calculation:** Calculate health indicator using Eq.(10) according to data measured on the monitored structural system. Then the final decision about the possibility of the monitored structure being healthy or damaged can be achieved.

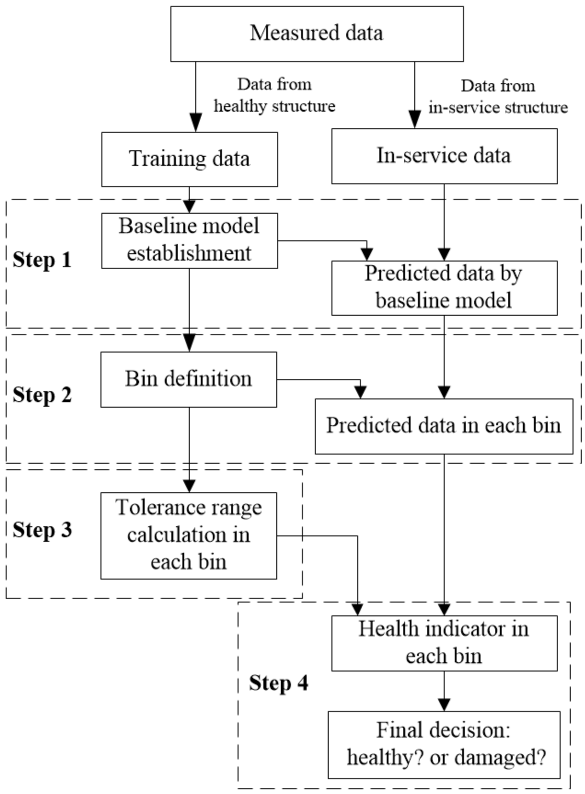


Fig. 8 Flowchart of the baseline model based SHM method

For the SHM method described above, the following remarks can be made regarding the measured data, baseline model, bins, tolerance range, and health indicator.

1) Measured data are involved in all steps. Measured data include both environment parameters and measurements which are sensitive to damage, for example, vibration, acoustic emission. Data involved in Step1-Step3 are measured from the structural system which is healthy and subject to no damage; while data involved in Step 4 are in-service data and measured from the monitored structural system. It should be pointed out that measured data involved in Step1-Step3 should cover all possible environmental conditions, or else SHM in that condition is limited.

2) Baseline model in Step 1 can represent the relationship between the healthy SHM feature and corresponding environment parameters. Therefore, the quality of

baseline model has a significant impact on eliminating the effect of varying environment. Knots, order of B-spline basis functions should be carefully chosen in order to obtain a high quality B-spline based baseline model.

3) Bins in Step 2 are divided according to environment parameters which means that the volume of each bin can be equal or unequal. But it is suggested that environmental conditions where SHM features have a similar damage sensitivity level are allocated in the same bin.

4) Both tolerance range in Step 3 and health indicator in Step 4 are statistical concepts. Therefore, massive data should be involved in both Step 3 and Step 4, tolerance range and health indicator are meaningless if only few data are involved. The threshold value for the health indicator to distinguish between damage and normal condition should be 1 under the ideal condition, but in practice, it is smaller than 1 due to many factors including modelling error, calculation error and measurement noise et al. The threshold value can be determined by the statistical analysis on the healthy condition. The threshold is a static for a particular structure because the influence of varying environment parameters has been considered in the baseline model.

3 Experimental case study

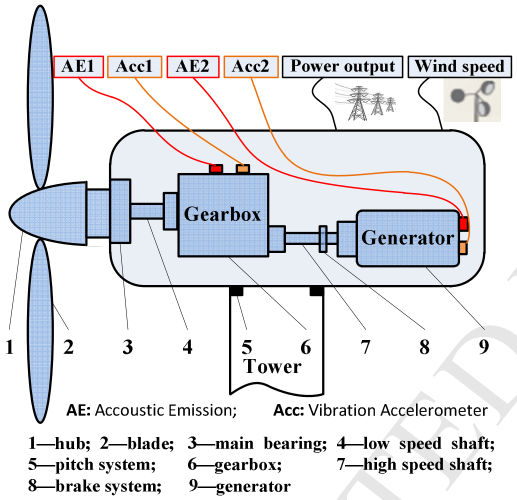
In order to demonstrate the ability of the proposed structural health monitoring method in practical applications, it is applied to monitor the health conditions of gearbox and generator in an operating wind turbine (WT) in this section.

3.1 Experimental measurements

Experimental measurements were undertaken in an operating wind turbine with type of 300KW-25 WINDMASTER located in the Wansbeck Blyth Harbour Wind Farm, UK. The major components of the monitored wind turbine are illustrated in Fig. 9. The function of gearbox is to transform input power from hub to high speed shaft, and the generator is to transmit mechanical power to electrical power. Thus, the gearbox and the generator are two of the most critical components for wind turbine; but gearbox, generator and corresponding shafts and bearings degrade slowly with operating time. Detection failures of such vital components are very important [24, 25, 32]. Therefore, the health conditions of gearbox and generator in the operating wind turbine are monitored in this experimental study.

In the experiment, two vibration accelerometers (Acc) and two acoustic emission(AE) sensors are mounted on the top of the gearbox (labelled as Acc1 and AE1) and at the back of the generator (labelled as Acc2and AE2) respectively as demonstrated in

Fig.9. The type of vibration accelerometers is B&K 8309, and the type of acoustic emission sensors is vallen VS 900RIC. Data from 4 sensors are recorded by the National Instruments (NI) data acquisition equipment with 4-Channel 20MHz simultaneous analogue input which is located at the bottom of tower and connected with sensors by a cable with length of 50 meters. Data were collected at different wind speeds discontinuously. During each data collection, one second data acquisition from the accelerometers and AE sensors were recorded as time driven data which can be considered as stationary signals. The sampling rate is 5M Hz. Meanwhile, the average values of the wind speeds and power outputs over a ten minutes period were also recorded which were considered as the representative of the environmental conditions, as shown in Fig. 10. Root Mean Square (RMS) of each sensor measurement for each data recording was treated as the damage-sensitive feature at the corresponding wind speed and power output which can be treated as hit driven data in this experimental case study.



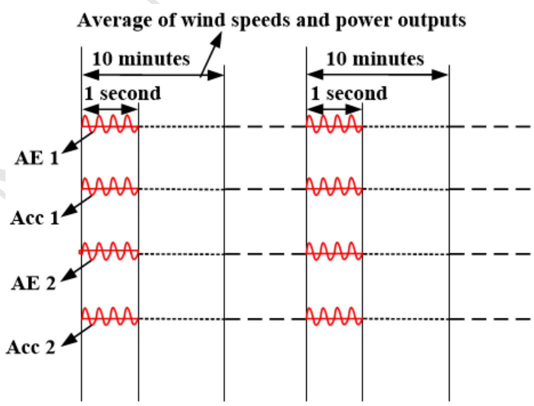


Fig.9 Main components of monitored WT

Fig. 10 Data acquisition schedule

The details of experiments are summarized in Table 3 where it can be observed that: two different state conditions of the wind turbine were investigated, one condition is no damage occurred in WT, the other condition is maintenance has been conducted before experiments. The Experiment #1 and #2 were conducted under the first condition while the Experiment #3 and #4 were conducted under the second condition. The data collected from Experiment #1 were used to obtain the improved B-spline based baseline model and the tolerance range of SHM features; the data collected from Experiment #2- #4 were used to prove the effectiveness of the proposed structural health method.

It should be pointed out that it is impossible to inject damage into healthy wind turbine systems without great expense, the measurements were conducted on an operating wind turbine without artificial damage. In order to solve this problem, apart from two experiments on the wind turbine without damage, another two experiments were conducted after maintenance and labelled as Experiment #3 and Experiment #4, the time interval of which was about two months, to verify the ability of the proposed method in distinguishing different healthy conditions.

Table 3 Details of the experiments

Experiments	State Condition Under Which Experiment Was Conducted		Usage of Data
Experiment #1	No damage	wind speed was from 4.7 to 24.8m/s; power output was from -15.9 to 302.7Kw	**Training data**: to obtain the improved B-spline based baseline model and the tolerance range of SHM features in each bin
Experiment #2		wind speed was from 5.0 to 24.0 m/s; power output was from -12.9 to 302.2Kw	**In-service data**: to prove effectiveness of the proposed SHM method when there was no damage in the system
Experiment #3	After maintenance	wind speed was from 5.5 to 19.5m/s; power output was from -15.0 to 302.0Kw	**In-service data**: to prove effectiveness of the proposed SHM method when the health condition of the system changed
Experiment #4		wind speed was from 5.0 to 15.3m/s; power output was from -15.5 to 251.7Kw	**In-service data**: to prove effectiveness of the proposed SHM method when the health condition of the system changed

3.2 Experimental data analysis

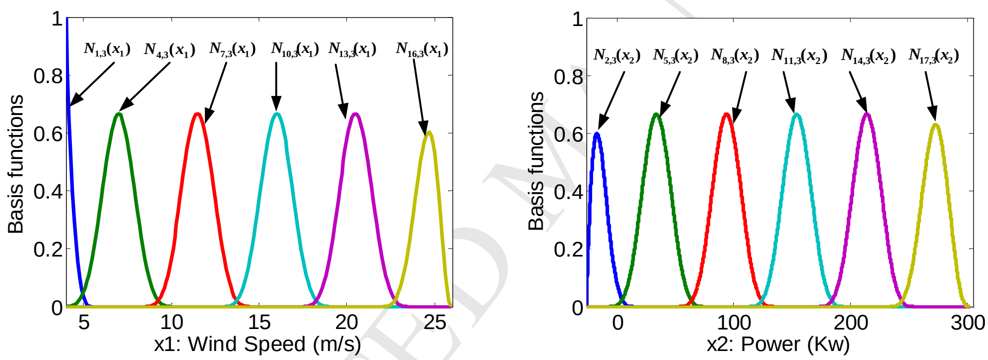
The results of the experimental study obtained at each step of the proposed method are given as follows.

Step 1: Baseline model establishment

The measured data from Experiment #1 are used to build the improved B-spline based baseline model by RFROE method in Eqs.(2)- (7). All data from experiment #1 are divided into 5 groups, the data in the first group are used to fit the improved B-spline based baseline model and the remaining ones are used to validate the baseline model by assessing the mean square error (MSE).

When wind speed is represented by x_1, power output is represented by x_2, and the order of basis functions is set as 3, the improved B-spline model for the relationship between the predicted signal feature y and x_1, x_2 can be derived from Eq. (2)-(7) . In this experimental case study, it is assumed that there are 16 knots for variable x_1 and 18 knots for variable x_2, then B-spline basis functions $N_{i_1,3}(x_1)$ and $N_{i_2,3}(x_2)$can be determined according to Eqs.(3.1) and(3.2), and some of them are shown in Fig. 11. By using the RFROE method in Eqs.(4)-(7), when error reduction ratios (ERRs) are set as 0.97, 0.93, 0.989, 0.975 for signals measured from AE1, AE 2, Acc1 and Acc2, respectively, the significant B-spline basis functions and corresponding coefficients are obtained. The first five selected terms and corresponding coefficients for each sensor measurement are listed in Table 4. Consequently, the baseline model is determined by the improved B-spline based model with B-spline basis functions, selected terms and corresponding coefficients.

The suitability of the obtained B-spline based baseline models is validated by assessing MSE with remaining 4 data groups which are not involved in the modelling process, the results are illustrated by bar charts in Fig. 12. Ideally, MSEs for the data groups not used in the modelling process are the same as that for modelling data, but because of inevitable modelling error and calculation error, MSEs for the data groups not used in the modelling process are always in the similar levels which are slightly higher than that for modelling data. It can be observed that the values of MSEs for the data groups not used in the modelling process are all slightly different from those for modelling data. So the modelling results are validated and therefore can be used for structural health monitoring.

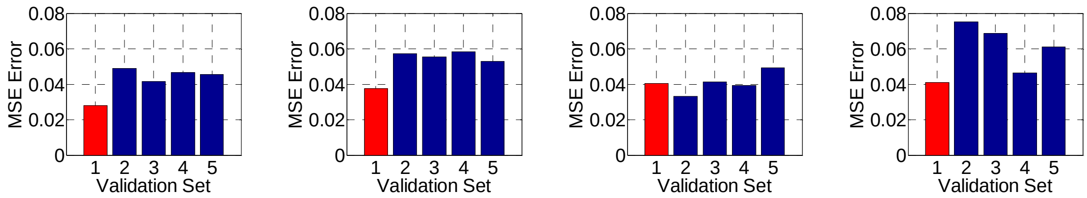


(a) Wind speed (b) Power

Fig. 11 Basis functions for improved B-spline model

Table 4 First five selected terms and corresponding coefficients

AE 1		AE 2		Acc 1		Acc 2	
Terms	α_{i_1,i_2}	terms	α_{i_1,i_2}	terms	α_{i_1,i_2}	terms	α_{i_1,i_2}
$\alpha_{0,0}$	0.0968	$\alpha_{0,0}$	0.0088	$\alpha_{0,0}$	2.2073	$\alpha_{0,0}$	1.9922
$N_{18,3}(x_2)$	0.2759	$N_{3,3}(x_2)$	-0.0145	$N_{3,3}(x_2)$	-3.0953	$N_{3,3}(x_2)$	-2.9660
$N_{16,3}(x_2)$	0.2276	$N_{10,3}(x_2)$	0.0168	$N_{18,3}(x_2)$	1.2158	$N_{6,3}(x_1)$	-0.9345
$N_{3,3}(x_2)$	-0.1659	$N_{18,3}(x_2)$	0.0190	$N_{16,3}(x_2)$	0.8161	$N_{18,3}(x_2)$	3.0551
$N_{6,3}(x_2)$	-0.0826	$N_{15,3}(x_1)N_{7,3}(x_2)$	0.0612	$N_{5,3}(x_2)$	-1.1536	$N_{16,3}(x_2)$	2.6837
$N_{14,3}(x_1)N_{7,3}(x_2)$	0.3021	$N_{8,3}(x_1)$	0.0092	$N_{17,3}(x_2)$	0.6098	$N_{5,3}(x_1)N_{3,3}(x_2)$	-6.2069



(a) Acoustic emission 1 (b) Acoustic emission 2 (c) Vibration 1 (d) Vibration 2

Fig. 12 Validation of each model

Step 2: Bin definition

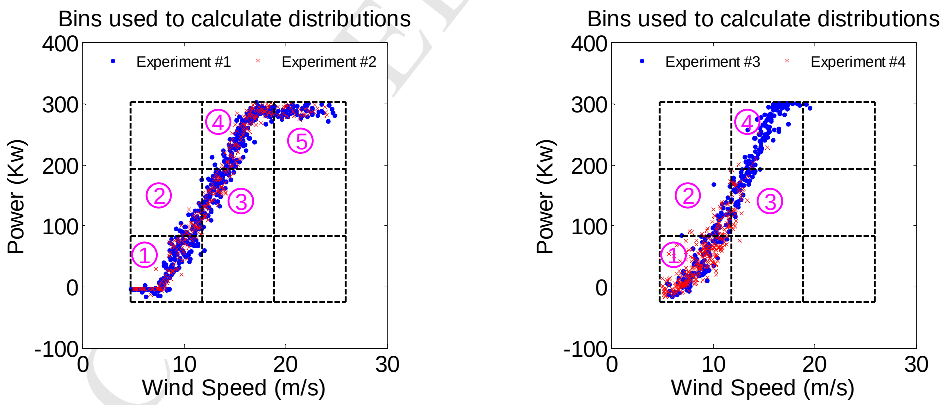
Bins are defined according to wind speeds and power outputs. When both wind speeds and power outputs are divided into three equal segments, the results are shown in Fig. 13. After neglecting bins where very few or no measured wind speeds and power outputs fall inside, 5 bins remain for Experiments #1 and #2, 4 bins for Experiment #3, and 3 bins for Experiment #4; all remaining bins are numbered as shown in Fig. 13.

Step 3: Tolerance range calculation

In each bin, the tolerance range of SHM features, which are RMS of measured signals in this study, is calculated separately using Eq.(9) according to data in Experiment #1.

Step 4: Health indicator calculation

Health indicator in each bin is calculated using Eq.(10) according to data in Experiments #2-#4, the results are shown in Table 5.



(a) Experiments #1-#2 (b) Experiments #3-#4

Fig. 13 Bins according to wind speeds and power outputs

Table 5 Health indicator for measurements in Experiments #2 - #4

Conditions	Experiment #2, No damage				Experiment #3, Maintenance				Experiment #4, Maintenance			
Location	AE1	AE2	Acc1	Acc2	AE1	AE2	Acc1	Acc2	AE1	AE2	Acc1	Acc2
Bin 1	0.988	0.988	1.000	0.988	0.960	0.901	0.396	0.713	0.949	0.864	0.670	0.777
Bin 2	1.000	1.000	1.000	1.000	0.971	0.371	0.600	0.914	1.000	0.404	0.173	0.981
Bin 3	1.000	0.964	1.000	1.000	0.889	0.044	0.800	0.933	1.000	0.345	0.276	0.966
Bin 4	0.977	0.989	1.000	0.955	0.838	0.045	0.955	0.991	--	--	--	--
Bin 5	1.000	1.000	1.000	0.906	--	--		--	--	--	--	--

Results analysis

It can be seen from Table 5 that the numbers of health indicators in Experiments #3-#4 are less than those in Experiment #2 because few data were collected in Experiments #3-#4 when wind speeds and power outputs were large as shown in Fig. 13; health indicators in different Bins are different which proves that changes of SHM features vary with the environmental conditions. In addition, for measurements in Experiment #2, health indicator in each bin is large, which indicates that both gearbox and generator are in good health condition. This indication is consistent with the practical situation of the wind turbine as stated in Table 3. For measurements in Experiment #3, some health indicators from AE sensor at the back of generator (AE2) and vibration accelerometer on the top of gearbox (Acc1) are small, which indicate that there are some changes in both gearbox and generator. The same conclusion can be reached by health indicators for measurements in Experiment #4. These are also consistent with the practical situation of the wind turbine as stated in Table 3. Therefore, the effectiveness of the proposed SHM method has been proved. However, health indicators for measurements from the AE sensor on the top of the gearbox (AE1) and vibration accelerometer at the back of the generator (Acc2) are large, indicating good health condition of both gearbox and generator. This means vibration is more sensitive to the condition change in the gearbox while AE signal is more sensitive to the condition variation in the generator. This conclusion is clearly very helpful for the choice of appropriate sensors for the health monitoring of various wind turbine components.

It should be pointed out that the application of the proposed technique is not limited to wind turbine gearbox/generator; it is feasible to many SHM applications particularly when the changes revealed by damage-sensitive features are affected by the working environment.

4 Conclusions

In this study, a baseline model based structural health monitoring method has been developed and its effectiveness has been investigated by experimental and simulation cases studies. Procedure with four steps is developed to guide how to implement the proposed structural health monitoring method. The analysis of the field data from an operating wind turbine has demonstrated that the new baseline model based structural health monitoring technique can distinguish different healthy conditions of gearbox and generator in WT. It can also be concluded from the field data analysis that vibration and AE signals are sensitive to condition changes of the gearbox and generator respectively, and the choosing sensor locations in experimental case study are applicable to the real industry.

5 Acknowledgment

The authors would like to express thanks to the National Natural Science Foundation of China (Grant Number: 11702318) and the Beijing Natural Science Foundation (Grant Number: 3184053) for their financial support of this research.

6 References

1. Farrar, C.R. and K. Worden, *An introduction to structural health monitoring.* Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences, 2007. **365**(1851): p. 303-315.

2. Malekjafarian, A., P.J. McGetrick, and E.J. OBrien, *A Review of Indirect Bridge Monitoring Using Passing Vehicles.* Shock and Vibration, 2015. **2015**: p. Article ID 286139, 16 pages.

3. Gao, Z.W., C. Cecati, and S.X. Ding, *A Survey of Fault Diagnosis and Fault-Tolerant Techniques-Part I: Fault Diagnosis With Model-Based and Signal-Based Approaches.* Ieee Transactions on Industrial Electronics, 2015. **62**(6): p. 3757-3767.

4. Gao, Z.W., C. Cecati, and S.X. Ding, *A Survey of Fault Diagnosis and Fault-Tolerant Techniques-Part II: Fault Diagnosis With Knowledge-Based and Hybrid/Active Approaches.* Ieee Transactions on Industrial Electronics, 2015. **62**(6): p. 3768-3774.

5. Ma, H., et al., *Fixed-point rubbing fault characteristic analysis of a rotor system based on contact theory.* Mechanical Systems and Signal Processing,

2013. **38**(1): p. 137-153.

6. Ma, H., et al., *Oil-film instability simulation in an overhung rotor system with flexible coupling misalignment.* Archive of Applied Mechanics, 2015. **85**(7): p. 893-907.

7. Ma, H., et al., *Analysis of dynamic characteristics for a rotor system with pedestal looseness.* Shock and Vibration, 2011. **18**(1-2): p. 13-27.

8. Kar, C. and A.R. Mohanty, *Monitoring gear vibrations through motor current signature analysis and wavelet transform.* Mechanical Systems and Signal Processing, 2006. **20**(1): p. 158-187.

9. Williams, T., et al., *Rolling element bearing diagnostics in run-to-failure lifetime testing.* Mechanical Systems and Signal Processing, 2001. **15**(5): p. 979-993.

10. Wang, S.Q., M. Zhang, and H.J. Li, *Damage Localization of an Offshore Platform considering Temperature Variations.* Mathematical Problems in Engineering, 2015. **2015**: p. Article ID 954926, 10 pages.

11. Sohn, H., *Effects of environmental and operational variability on structural health monitoring.* Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences, 2007. **365**(1851): p. 539-560.

12. Sohn, H., et al. *Online Damage Detection for Theme Park Rides*. in *Proceedings of 22nd International Modal Analysis Conferenc*. 2004. Dearborn, MI.

13. Ha, T.M., S. Fukada, and K. Torii, *Long-Term Vibration Monitoring of the Effects of Temperature and Humidity on PC Girders with and without Fly Ash considering ASR Deterioration.* Shock and Vibration, 2017. **2017**: p. Article ID 5468950, 23 pages.

14. Han, Q.K., et al., *Periodic Motion Stability of a Dual-Disk Rotor System with Rub-Impact at Fixed Limiter.* Vibro-Impact Dynamics of Ocean Systems and Related Problems, 2009. **44**: p. 105-119.

15. Loutas, T.H., et al., *The combined use of vibration, acoustic emission and oil debris on-line monitoring towards a more effective condition monitoring of rotating machinery.* Mechanical Systems and Signal Processing, 2011. **25**(4): p. 1339-1352.

16. Zolna, K., et al., *Nonlinear Cointegration Approach for Condition Monitoring of Wind Turbines.* Mathematical Problems in Engineering, 2015. **2015**: p. Article ID 978156, 11 pages.

17. Surace, C. and K. Worden, *Novelty detection in a changing environment: A negative selection approach.* Mechanical Systems and Signal Processing, 2010. **24**(4): p. 1114-1128.

18. ZHao, X., *New Methods for Structural Health Monitoring and Damage Localization* in *Department of Automatic Control and Systems Engineering,* 2015, University of Sheffield.

19. Makis, V. and M. Yang, *ARX model-based gearbox fault detection and localization under varying load conditions.* Journal of Sound and Vibration, 2010. **329**(24): p. 5209-5221.

20. Worden, K., H. Sohn, and C.R. Farrar, *Novelty detection in a changing environment: Regression and interpolation approaches.* Journal of Sound and Vibration, 2002. **258**(4): p. 741-761.

21. Zhao, X. and Z. Lang, *A novel health probability for structural health monitoring*, in *Proceedings of the 18th International Conference on Automation & Computing*2012: Loughborough University, Leicestershire, UK.

22. Cross, E.J., K. Worden, and Q. Chen, *Cointegration: a novel approach for the removal of environmental trends in structural health monitoring data.* Proceedings of the Royal Society a-Mathematical Physical and Engineering Sciences, 2011. **467**(2133): p. 2712-2732.

23. Cross, E.J., et al., *Features for damage detection with insensitivity to environmental and operational variations.* Proceedings of the Royal Society a-Mathematical Physical and Engineering Sciences, 2012. **468**(2148): p. 4098-4122.

24. Elforjani, M., S. Shanbr, and E. Bechhoefer, *Detection of faulty high speed wind turbine bearing using signal intensity estimator technique.* Wind Energy, 2018. **21**(1): p. 53-69.

25. Elforjani, M. and E. Bechhoefer, *Analysis of extremely modulated faulty wind turbine data using spectral kurtosis and signal intensity estimator.* Renewable Energy, 2018. **127**: p. 258-268.

26. Hills, A.F., et al., *A novel baseline model-based technique for condition monitoring of wind turbine components.* The British Insitute of Non-Destructive Testing, 2011. **53**(8): p. 434-438.

27. Sohn, H., K. Worden, and C.R. Farrar, *Statistical damage classification under changing environmental and operational conditions.* Journal of Intelligent Material Systems and Structures, 2002. **13**(9): p. 561-574.

28. Prautzsch, H., W. Boehm, and M. Paluszny, *Bézier and B-spline techniques*.

2010, Berlin ; New York: Springer.

29. Han, Q., Z. Zhang, and B. Wen, *Periodic motions of a dual-disc rotor system with rub-impact at fixed limiter.* Proceedings of the Institution of Mechanical Engineers Part C-Journal of Mechanical Engineering Science, 2008. **222**(10): p. 1935-1946.

30. Billings, S.A., S. Chen, and M.J. Korenberg, *Identification of Mimo Non-Linear Systems Using a Forward-Regression Orthogonal Estimator.* International Journal of Control, 1989. **49**(6): p. 2157-2189.

31. Billings, S.A. and H.L. Wei, *An adaptive orthogonal search algorithm for model subset selection and non-linear system identification.* International Journal of Control, 2008. **81**(5): p. 714-724.

32. Marquez, F.P.G., et al., *Condition monitoring of wind turbines: Techniques and methods.* Renewable Energy, 2012. **46**: p. 169-178.

Highlights

- Effects of varying environment are considered when conducting structural health monitoring.
- Ability of the proposed method is verified by monitoring the health conditions of gearbox and generator in an operating wind turbine.
- Proposed method can be applied for condition monitoring of other structures and components.
- Choice of appropriate sensors for health monitoring of various wind turbine components is concluded.