



This is a repository copy of *Baseline model based structural health monitoring method under varying environment*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/144539/>

Version: Accepted Version

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>

Article available under the terms of the CC-BY-NC-ND licence
(<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

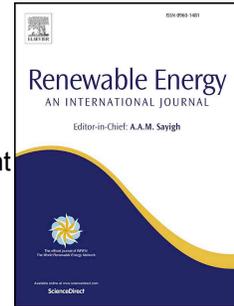


eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

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

Please cite this article as: Zhao X, Lang Z, Baseline model based structural health monitoring method under varying environment, *Renewable Energy* (2019), doi: <https://doi.org/10.1016/j.renene.2019.02.007>.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

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

47 However, the changes revealed by damage-sensitive features which are always
48 considered as SHM features are affected not only by damage in the inspected
49 structural system but also by the working environment [10]. The varying environment
50 has significant impacts on the system dynamic behaviours as discussed by Sohn in
51 [11]. Moreover, Sohn et al. [12] studied the vibration of a theme park ride by
52 combining time series analysis with statistical pattern recognition technique and
53 concluded that the feature variation caused by mass loading was more obvious than
54 that caused by delamination damage. Ha et al. [13] researched the effects of
55 temperature and humidity on pre-stressed concrete girders and found that when the
56 temperature and humidity increased, the frequencies and damping ratios decreased
57 in proportion. The stability of a rotor system with rub-impact damage under different
58 rotating speeds was investigated by Han et al. in [14], and the results revealed that
59 when rotating speed increased, the system exhibited firstly stable, then
60 period-doubling bifurcations and finally reached the stable periodic motion again. As
61 for the gearbox, Loutas et al. [15] researched how the features of the vibration and
62 AE signals in the frequency domain changed when the gearbox kept working until
63 several teeth were cut and considerable damage happened on the shaft. It was
64 concluded that the oil temperature had an effect on the recordings.

65 Therefore, many researchers have paid attention to the influence of varying
66 environment on system behaviors, and then, try to investigate the effect of
67 non-damage factors so as to enhance the reliability of structural health monitoring
68 methods[16-18]. One type of methods for removing the effects of varying
69 environment is to model the relationship of damage-sensitive features and varying
70 environment. Makis and Yang [19] found that a model developed under the constant
71 load assumption could not recognize whether the vibration feature changes of
72 gearbox were caused by the load variation or by a failure occurrence. To settle this
73 problem, an ARX model was proposed which considered load as additional
74 information. Worden et al. [20] revised the conventional outlier analysis method by

75 replacing the traditional mean vector of damage-sensitive features with features at
76 the same temperature predicted from a polynomial regression model in temperature
77 and the mean vector of damage-sensitive features at this temperature. Zhao and
78 Lang established the relationship between the varying environments and SHM
79 features using a polynomial model [21] and a B-spline model [18] respectively, and
80 then proposed a novel health indicator after removing the environmental effect to
81 indicate health condition of the monitored system. Experimental study on wind
82 turbine components proved the effectiveness of the health indicator. Another type of
83 methods removing the effects of varying environment is to extract signal features
84 which are insensitive to environmental variation but still damage-sensitive. Cross and
85 Worden combined linearly several damage-sensitive features to produce a new
86 feature which was independent of environmental variation but was sensitive to
87 damage in [22], and further tried to extract signal features which were insensitive to
88 environmental variation but still damage-sensitive by co-integration technique,
89 outlier analysis and minor principal components techniques in [23].

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

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

109 **2 Methodology**

110 Traditionally, structural health monitoring is achieved by monitoring structural signal
111 features and identifying any deviation of these features from a healthy one, an
112 obvious deviation is indicative of a developing damage. The signal feature of the

113 monitored structure can be named as in-service feature, while the signal feature of
 114 the healthy structure can be named as health feature. They are extracted
 115 respectively from sensor measurements of the monitored structure and the health
 116 structure by using a range of data analysis methods [11], such as time domain
 117 analysis, frequency domain analysis or time-frequency domain analysis [9, 24, 25].
 118 However, fluctuating environment has significant impacts on the structure
 119 performance, and can also cause the change of signal features which will lead to
 120 incorrect results of SHM. In order to remove the effect of fluctuating environment on
 121 the results of traditional structural health monitoring, a baseline model is proposed
 122 to represent the relationship between healthy SHM features and corresponding
 123 environment parameters. Then tolerance ranges of the in-service SHM features
 124 under certain environment conditions are obtained by statistical analysis. Finally,
 125 in-service structural system health condition can be determined by identifying
 126 occurrences of in-service SHM features within tolerance range. Baseline model,
 127 tolerance range and health indicator are achieved as follows.

128 2.1 B-spline based baseline model

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

$$135 \quad y = f(x_1, x_2, x_3, \dots, x_M) \quad (1)$$

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

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

$$142 \quad y = f(x_1, x_2, x_3, \dots, x_M) = \sum_{i_1=0}^{M_1} \dots \sum_{i_M=0}^{M_M} \alpha_{i_1, i_2, \dots, i_M} N_{i_1, p}(x_1) \dots N_{i_M, p}(x_M) \quad (2)$$

143 where $N_{i_1, p}(x_1), \dots, N_{i_M, p}(x_M)$ are the $i_1^{\text{th}}, \dots, i_M^{\text{th}}$ B-spline basis functions of degree
 144 p with respect to variables x_1, \dots, x_M , respectively; and $N_{i_1, p}(x_1), \dots, N_{i_M, p}(x_M)$ can
 145 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
 146 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
 147 $N_{i_m, p}(x_m)$, where $m = 1, 2, \dots, M$. Given a knot vector $\mathbf{x}_m = \{x_{m,0}, x_{m,1}, x_{m,2}, \dots, x_{m,K}\}$

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

$$150 \quad N_{i_m,0}(x_m) = \begin{cases} 1 & \text{if } x_{m,i_m} \leq x_m < x_{m,i_m+1} \\ 0 & \text{otherwise} \end{cases} \quad (3.1)$$

$$151 \quad N_{i_m,p}(x_m) = \frac{x_m - x_{m,i_m}}{x_{m,i_m+p} - x_{m,i_m}} N_{i_m,p-1}(x_m) + \frac{x_{m,i_m+p+1} - x_m}{x_{m,i_m+p+1} - x_{m,i_m+1}} N_{i_m+1,p-1}(x_m) \quad (3.2)$$

152 It can be deduced from Eqs.(2) and (3) that the basis function $N_{i_m,p}(x_m)$ is non-zero
 153 on only $p + 1$ knot spans, namely, $[x_{m,i_m}, x_{m,i_m+1})$, $[x_{m,i_m+1}, x_{m,i_m+2})$, ...,
 154 $[x_{m,i_m+p}, x_{m,i_m+p+1})$, and on any knot span $[x_{m,i_m}, x_{m,i_m+1})$, at most $p + 1$ basis
 155 functions with degree p are non-zero, namely, $N_{i_m-p,p}(x_m)$, $N_{i_m-p+1,p}(x_m)$, ...,
 156 $N_{i_m,p}(x_m)$. Thus, changing the control coefficient $\alpha_{i_1, i_2, \dots, i_M}$ or the position of knot
 157 $x_{m,i}$ only affects the curve shape of B spline model on local span; this is so-called
 158 local effect or local modification property. In addition, B-spline curve expressed by
 159 Eq.(3) is a piecewise and derivative curve with each component a curve of degree p ,
 160 this property allows B-spline model to fit complex shapes smoothly with lower
 161 degree than ARX model and with higher efficiency than neural network. The B-spline
 162 model expressed by Eqs.(2) and (3) has excellent capabilities in smooth data fitting
 163 and local effect, and can be used to fit the data with lower degree but higher
 164 efficiency, so it is employed in this paper to determine the baseline model.

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

$$x_1 = \{x_{1,0}, \dots, x_{1,K}\} = \{60, 61, 63.3, 66.6, 69.9, 73.2, 76.5, 79.8, 83.1, 86.4, 89.7, 93, 96.3, 99.6, 100\}$$

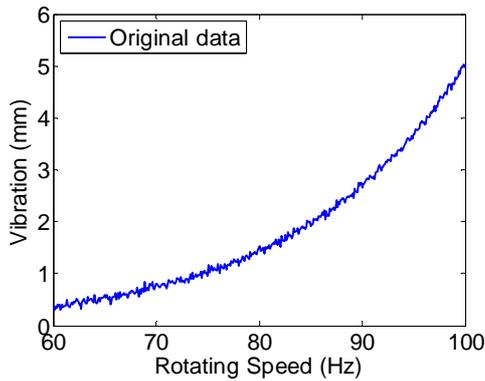
174 Then, B-spline basis functions $N_{i_1,p}(x_1)$ can be determined according to Eqs.(3.1)
 175 and (3.2), and some of them are shown in Fig. 2. Corresponding coefficients are
 176 estimated by least square, and the results are listed in Table 1. The fitting error when
 177 $K = 14$ is shown in Fig. 3 by a blue solid line. The maximum, mean and standard
 178 deviation are 0.1636, 0.0032 and 0.0582 respectively, indicating that the fitting error
 179 is small and ignorable. Therefore, data in Fig.1 can be represented by B-spline model
 180 Eqs.(2)-(3) with B-spline basis functions in Fig.2 and corresponding coefficients in
 181 Table 1.

182

183

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



184

185

Fig.1 Original data

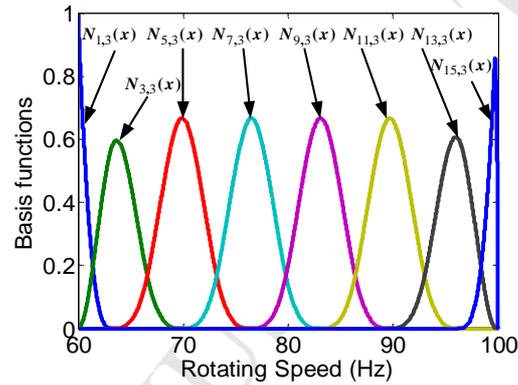
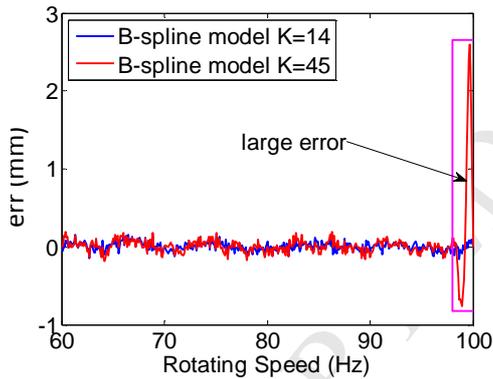


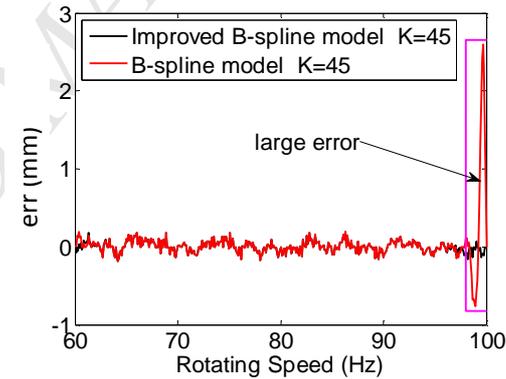
Fig. 2 B-spline basis functions



186

187

Fig. 3 Fitting error by B-spline model Fig. 4 Fitting error by improved B-spline model



188 However, when the number of knots K increases, the performance of B-spline
 189 model in fitting data becomes unstable, e.g., when the number of knots increases
 190 to $K = 45$, fitting error by B-spline model at the end data is much larger than that
 191 when $K = 14$ as shown in Fig.3; the maximum, mean and standard deviation are
 192 2.5892, 0.0284 and 0.2725 respectively. This is because corrosion of data at the end
 193 tends to deteriorate when the number of knots and B-spline basis functions become
 194 larger. Besides, the increase in the number of knots and B-spline basis functions will
 195 also lead to more complicated and tedious computations, and computational errors
 196 are accumulated when fitting a B-spline model. In order to solve this problem,
 197 conventional B-spline model can be improved by reordering all remaining B-spline
 198 basis functions and/or ignoring insignificant B-spline basis functions and their
 199 multiplications by using recursive forward-regression orthogonal estimator (RFROE)

200 [30]. The terms which contribute prominently to the model can be selected as
201 follows.

202 **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
203 considered as possible candidates for the most important term $w_1(t)$. For
204 $i_1 = 0,1,2, \dots, M_1, \dots, i_M = 0,1,2, \dots, M_M$, set $w_1^{(i_1 \dots i_M)}(t) = N_{i_1,p}(x_1) \dots N_{i_M,p}(x_M)$, then
205 calculate

$$206 \quad \hat{g}_1^{(i_1 \dots i_M)} = \frac{\sum_{t=1}^N w_1^{(i_1 \dots i_M)}(t) y(t)}{\sum_{t=1}^N \left(w_1^{(i_1 \dots i_M)}(t) \right)^2} \quad (4)$$

207 and

$$208 \quad [err]_1^{(i_1 \dots i_M)} = \frac{\left(\hat{g}_1^{(i_1 \dots i_M)} \right)^2 \sum_{t=1}^N \left(w_1^{(i_1 \dots i_M)}(t) \right)^2}{\sum_{t=1}^N y^2(t)} \quad (5)$$

209 **Step (b):** Find the maximum of $[err]_1^{(i_1 \dots i_M)}$, e.g., $[err]_1^{(M_1 \dots M_M)} = \max\{[err]_1^{(i_1 \dots i_M)}, i_1 =$
210 $0,1,2, \dots, M_1, \dots, i_M = 0,1,2, \dots, M_M\}$. Then the first term is selected with $[err]_1 = [err]_1^{(M_1 \dots M_M)}$,
211 and $w_1(t) = w_1^{(M_1 \dots M_M)}(t) = N_{M_1,p}(x_1) \dots N_{M_M,p}(x_M)$.

212 **Step (c):** All the remaining terms are considered as possible candidates for $w_2(t)$.
213 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
214 using Eqs.(4) and (5), respectively, where

$$215 \quad \alpha_{12}^{(i_1 \dots i_M)} = \frac{\sum_{t=1}^N w_1(t) N_{i_1,p}(x_1) \dots N_{i_M,p}(x_M)}{\sum_{t=1}^N w_1^2(t)} \quad (6)$$

216 **Step (d):** Find the maximum of $[err]_2^{(i_1 \dots i_M)}$, and then corresponding term
217 $N_{i_1,p}(x_1) \dots N_{i_M,p}(x_M)$ is selected.

218 **Step (e):** Then Step (c) and (d) are iterative, and the procedure is terminated at the
219 D_s^{th} step when

$$220 \quad 1 - \sum_{i=1}^{D_s} [err]_i < a \text{ desired tolerance}, D_s < D \quad (7)$$

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

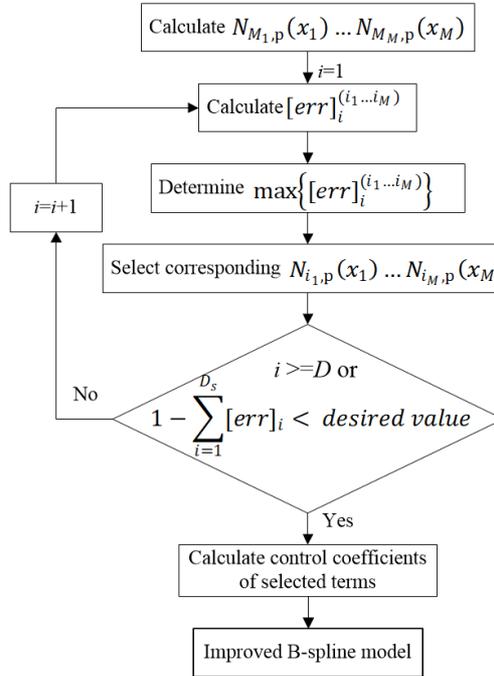
222 The value of the desired tolerance can be determined by using APRESS criteria in
223 [31].

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

226 The fitting error by using improved B-spline model method is shown in Fig.4. The

227 maximum, mean and standard deviation of the fitting error are 0.1885, 0.0009 and
 228 0.0708 respectively, indicating that the value of the fitting error by using improved
 229 B-spline model is obviously smaller than that by using conventional B-spline model.

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



232

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

234 2.2 Tolerance range

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

$$237 \quad e = y' - y \quad (8)$$

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

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

$$253 \quad y' = y + e \in [y + \mu - 3\sigma, y + \mu + 3\sigma] \quad (9)$$

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

257 2.3 Health indicator

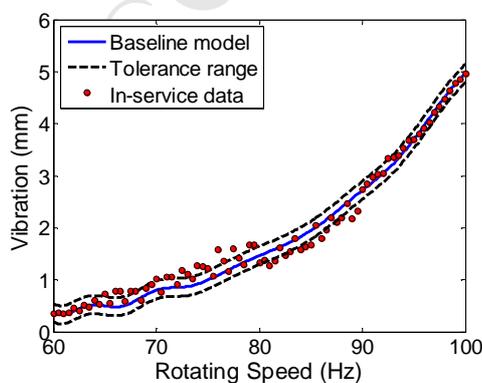
258 According to the definition of tolerance range above, if a monitored structure is
 259 operating in a healthy condition, most in-service y' should fall into the tolerance
 260 range. If there is a change or damage, only a small number of values of y' are
 261 within the corresponding tolerance range.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

$$263 \quad P = N_{in}/N_{all} \quad (10)$$

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

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

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



272

273 Fig. 6 Tolerance range and in-service data

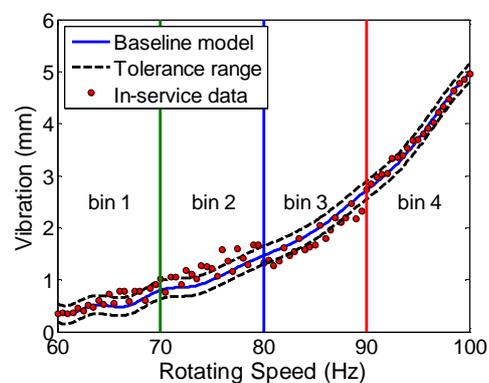


Fig. 7 Bins

274 2.4 Health indicator in each bin

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

$$283 \quad B_{n_1 n_2, \dots, n_M} = \{x_1, x_2, \dots, x_M\}, x_i \in [x_{i, n_i}, x_{i, n_i+1}] \quad (11)$$

284 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}
 285 are two edges of n_i^{th} segments for variable x_i ; $n_i = 1, 2, \dots, M_i$; M_i is the total
 286 number of the segments for i^{th} variable x_i . In order to describe bins more precisely,
 287 the bins are renumbered by the single subscript.

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

295 Table 2 Calculation of health indicator in each bin

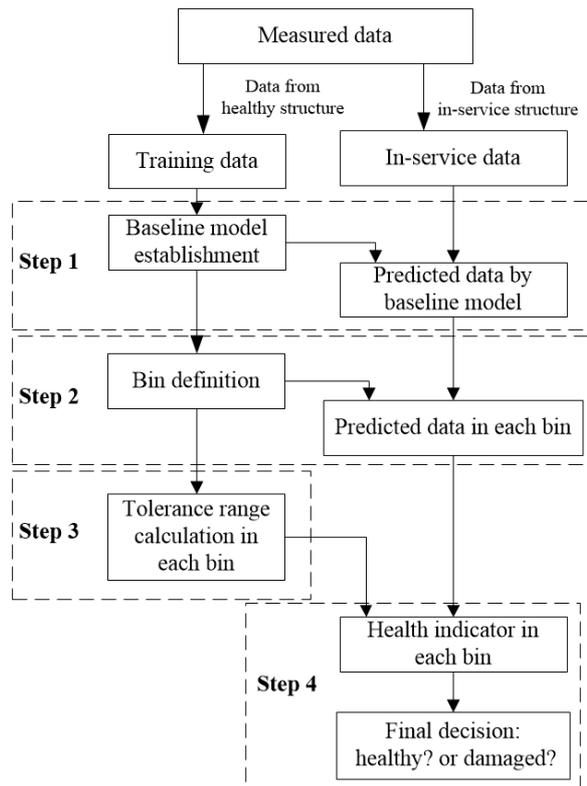
Bin index	N_{in}	N_{all}	P	Bin index	N_{in}	N_{all}	P
Bin 1	16	20	0.80	Bin 3	7	20	0.35
Bin 2	7	20	0.35	Bin 4	20	21	0.9524

296 2.5 Baseline model based SHM method and remarks

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

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

304 **Step 2: Bin definition:** Define bins using Eq.(11) according to the value of
 305 environment parameters;
 306 **Step 3: Tolerance range calculation:** Calculate tolerance range of SHM feature in
 307 each bin using Eq.(9) according to data measured on the healthy structural
 308 system;
 309 **Step 4: Health indicator calculation:** Calculate health indicator using Eq.(10)
 310 according to data measured on the monitored structural system. Then the
 311 final decision about the possibility of the monitored structure being healthy
 312 or damaged can be achieved.



313

314

Fig. 8 Flowchart of the baseline model based SHM method

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

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

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

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

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

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

343 **3 Experimental case study**

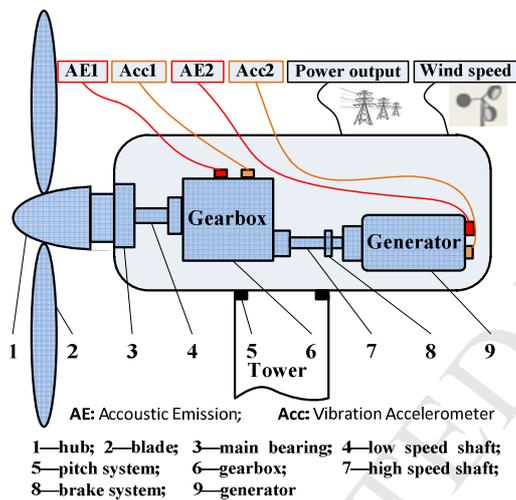
344 In order to demonstrate the ability of the proposed structural health monitoring
345 method in practical applications, it is applied to monitor the health conditions of
346 gearbox and generator in an operating wind turbine (WT) in this section.

347 **3.1 Experimental measurements**

348 Experimental measurements were undertaken in an operating wind turbine with
349 type of 300KW-25 WINDMASTER located in the Wansbeck Blyth Harbour Wind Farm,
350 UK. The major components of the monitored wind turbine are illustrated in Fig. 9.
351 The function of gearbox is to transform input power from hub to high speed shaft,
352 and the generator is to transmit mechanical power to electrical power. Thus, the
353 gearbox and the generator are two of the most critical components for wind turbine;
354 but gearbox, generator and corresponding shafts and bearings degrade slowly with
355 operating time. Detection failures of such vital components are very important [24,
356 25, 32]. Therefore, the health conditions of gearbox and generator in the operating
357 wind turbine are monitored in this experimental study.

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

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



375

376 Fig.9 Main components of monitored WT

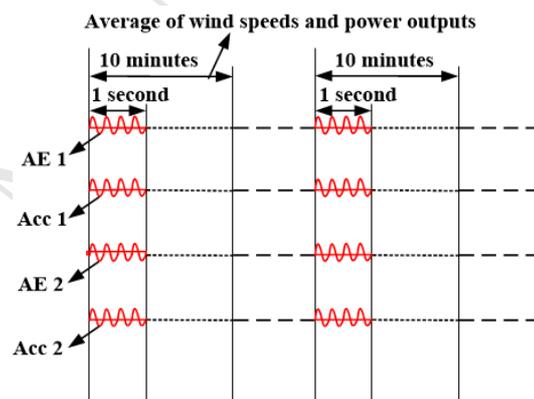


Fig. 10 Data acquisition schedule

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

386 It should be pointed out that it is impossible to inject damage into healthy wind
 387 turbine systems without great expense, the measurements were conducted on an
 388 operating wind turbine without artificial damage. In order to solve this problem,
 389 apart from two experiments on the wind turbine without damage, another two

390 experiments were conducted after maintenance and labelled as Experiment #3 and
 391 Experiment #4, the time interval of which was about two months, to verify the ability
 392 of the proposed method in distinguishing different healthy conditions.

393 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

394 3.2 Experimental data analysis

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

397 Step 1: Baseline model establishment

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

403 When wind speed is represented by x_1 , power output is represented by x_2 , and the
 404 order of basis functions is set as 3, the improved B-spline model for the relationship
 405 between the predicted signal feature y and x_1, x_2 can be derived from Eq. (2)-(7) .
 406 In this experimental case study, it is assumed that there are 16 knots for variable x_1
 407 and 18 knots for variable x_2 , then B-spline basis functions $N_{i,3}(x_1)$ and
 408 $N_{i,3}(x_2)$ can be determined according to Eqs.(3.1) and(3.2), and some of them are
 409 shown in Fig. 11. By using the RFROE method in Eqs.(4)-(7), when error reduction

410 ratios (ERRs) are set as 0.97, 0.93, 0.989, 0.975 for signals measured from AE1, AE 2,
 411 Acc1 and Acc2, respectively, the significant B-spline basis functions and
 412 corresponding coefficients are obtained. The first five selected terms and
 413 corresponding coefficients for each sensor measurement are listed in Table 4.
 414 Consequently, the baseline model is determined by the improved B-spline based
 415 model with B-spline basis functions, selected terms and corresponding coefficients.

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

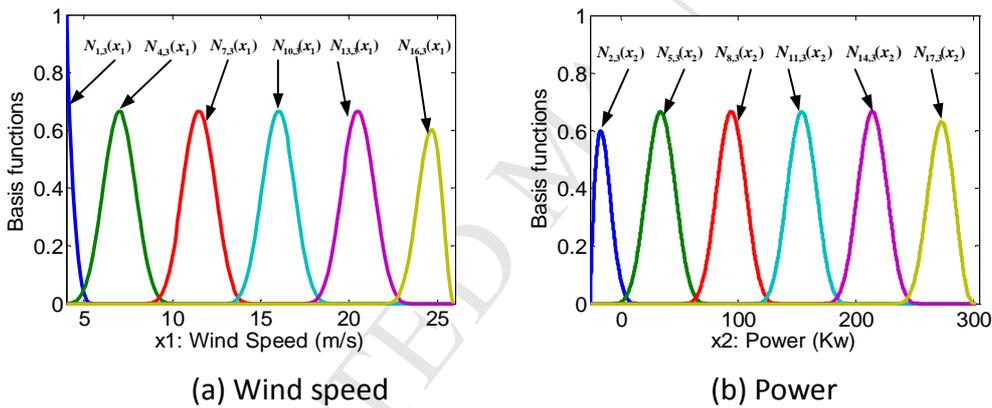
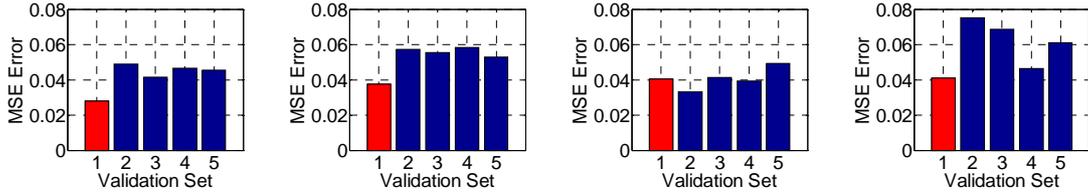


Fig. 11 Basis functions for improved B-spline model

429 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



430

431

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

432

Fig. 12 Validation of each model

433

Step 2: Bin definition

434

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.

440

Step 3: Tolerance range calculation

441

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.

442

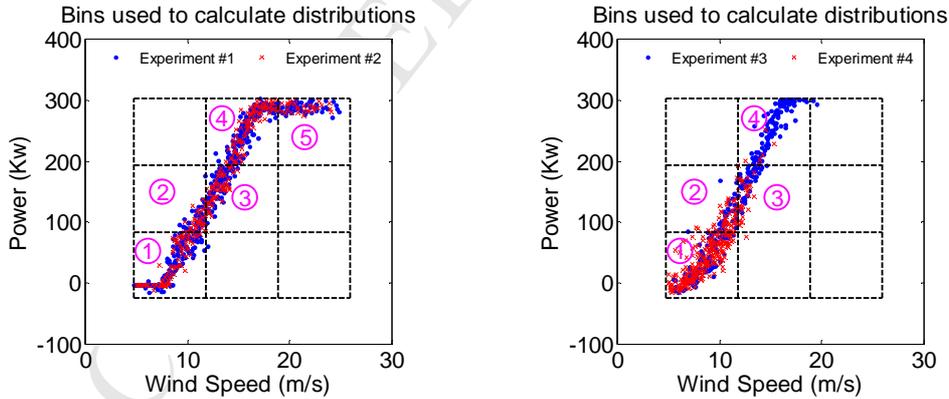
443

Step 4: Health indicator calculation

444

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

445



446

(a) Experiments #1-#2

(b) Experiments #3-#4

447

448

Fig. 13 Bins according to wind speeds and power outputs

449

450

451

452

453

454

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

Conditions	Experiment #2, No damage				Experiment #3, Maintenance				Experiment #4, Maintenance			
	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	--	--	--	--	--	--	--	--

455 Results analysis

456 It can be seen from Table 5 that the numbers of health indicators in Experiments
457 #3-#4 are less than those in Experiment #2 because few data were collected in
458 Experiments #3-#4 when wind speeds and power outputs were large as shown in Fig.
459 13; health indicators in different Bins are different which proves that changes of SHM
460 features vary with the environmental conditions. In addition, for measurements in
461 Experiment #2, health indicator in each bin is large, which indicates that both
462 gearbox and generator are in good health condition. This indication is consistent with
463 the practical situation of the wind turbine as stated in Table 3. For measurements in
464 Experiment #3, some health indicators from AE sensor at the back of generator (AE2)
465 and vibration accelerometer on the top of gearbox (Acc1) are small, which indicate
466 that there are some changes in both gearbox and generator. The same conclusion
467 can be reached by health indicators for measurements in Experiment #4. These are
468 also consistent with the practical situation of the wind turbine as stated in Table 3.
469 Therefore, the effectiveness of the proposed SHM method has been proved.
470 However, health indicators for measurements from the AE sensor on the top of the
471 gearbox (AE1) and vibration accelerometer at the back of the generator (Acc2) are
472 large, indicating good health condition of both gearbox and generator. This means
473 vibration is more sensitive to the condition change in the gearbox while AE signal is
474 more sensitive to the condition variation in the generator. This conclusion is clearly
475 very helpful for the choice of appropriate sensors for the health monitoring of
476 various wind turbine components.

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

481 **4 Conclusions**

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

492 **5 Acknowledgment**

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

496 **6 References**

- 497 1. Farrar, C.R. and K. Worden, *An introduction to structural health monitoring*.
498 Philosophical Transactions of the Royal Society a-Mathematical Physical and
499 Engineering Sciences, 2007. **365**(1851): p. 303-315.
- 500 2. Malekjafarian, A., P.J. McGetrick, and E.J. OBrien, *A Review of Indirect Bridge*
501 *Monitoring Using Passing Vehicles*. Shock and Vibration, 2015. **2015**: p. Article
502 ID 286139, 16 pages.
- 503 3. Gao, Z.W., C. Cecati, and S.X. Ding, *A Survey of Fault Diagnosis and*
504 *Fault-Tolerant Techniques-Part I: Fault Diagnosis With Model-Based and*
505 *Signal-Based Approaches*. IEEE Transactions on Industrial Electronics, 2015.
506 **62**(6): p. 3757-3767.
- 507 4. Gao, Z.W., C. Cecati, and S.X. Ding, *A Survey of Fault Diagnosis and*
508 *Fault-Tolerant Techniques-Part II: Fault Diagnosis With Knowledge-Based and*
509 *Hybrid/Active Approaches*. IEEE Transactions on Industrial Electronics, 2015.
510 **62**(6): p. 3768-3774.
- 511 5. Ma, H., et al., *Fixed-point rubbing fault characteristic analysis of a rotor*
512 *system based on contact theory*. Mechanical Systems and Signal Processing,

- 513 2013. **38**(1): p. 137-153.
- 514 6. Ma, H., et al., *Oil-film instability simulation in an overhung rotor system with*
515 *flexible coupling misalignment*. Archive of Applied Mechanics, 2015. **85**(7): p.
516 893-907.
- 517 7. Ma, H., et al., *Analysis of dynamic characteristics for a rotor system with*
518 *pedestal looseness*. Shock and Vibration, 2011. **18**(1-2): p. 13-27.
- 519 8. Kar, C. and A.R. Mohanty, *Monitoring gear vibrations through motor current*
520 *signature analysis and wavelet transform*. Mechanical Systems and Signal
521 Processing, 2006. **20**(1): p. 158-187.
- 522 9. Williams, T., et al., *Rolling element bearing diagnostics in run-to-failure*
523 *lifetime testing*. Mechanical Systems and Signal Processing, 2001. **15**(5): p.
524 979-993.
- 525 10. Wang, S.Q., M. Zhang, and H.J. Li, *Damage Localization of an Offshore*
526 *Platform considering Temperature Variations*. Mathematical Problems in
527 Engineering, 2015. **2015**: p. Article ID 954926, 10 pages.
- 528 11. Sohn, H., *Effects of environmental and operational variability on structural*
529 *health monitoring*. Philosophical Transactions of the Royal Society
530 a-Mathematical Physical and Engineering Sciences, 2007. **365**(1851): p.
531 539-560.
- 532 12. Sohn, H., et al. *Online Damage Detection for Theme Park Rides*. in
533 *Proceedings of 22nd International Modal Analysis Conferenc*. 2004. Dearborn,
534 MI.
- 535 13. Ha, T.M., S. Fukada, and K. Torii, *Long-Term Vibration Monitoring of the*
536 *Effects of Temperature and Humidity on PC Girders with and without Fly Ash*
537 *considering ASR Deterioration*. Shock and Vibration, 2017. **2017**: p. Article ID
538 5468950, 23 pages.
- 539 14. Han, Q.K., et al., *Periodic Motion Stability of a Dual-Disk Rotor System with*
540 *Rub-Impact at Fixed Limiter*. Vibro-Impact Dynamics of Ocean Systems and
541 Related Problems, 2009. **44**: p. 105-119.
- 542 15. Loutas, T.H., et al., *The combined use of vibration, acoustic emission and oil*
543 *debris on-line monitoring towards a more effective condition monitoring of*
544 *rotating machinery*. Mechanical Systems and Signal Processing, 2011. **25**(4): p.
545 1339-1352.
- 546 16. Zolna, K., et al., *Nonlinear Cointegration Approach for Condition Monitoring of*
547 *Wind Turbines*. Mathematical Problems in Engineering, 2015. **2015**: p. Article
548 ID 978156, 11 pages.

- 549 17. Surace, C. and K. Worden, *Novelty detection in a changing environment: A*
550 *negative selection approach*. Mechanical Systems and Signal Processing, 2010.
551 **24**(4): p. 1114-1128.
- 552 18. ZHao, X., *New Methods for Structural Health Monitoring and Damage*
553 *Localization in Department of Automatic Control and Systems Engineering*,
554 2015, University of Sheffield.
- 555 19. Makis, V. and M. Yang, *ARX model-based gearbox fault detection and*
556 *localization under varying load conditions*. Journal of Sound and Vibration,
557 2010. **329**(24): p. 5209-5221.
- 558 20. Worden, K., H. Sohn, and C.R. Farrar, *Novelty detection in a changing*
559 *environment: Regression and interpolation approaches*. Journal of Sound and
560 Vibration, 2002. **258**(4): p. 741-761.
- 561 21. Zhao, X. and Z. Lang, *A novel health probability for structural health*
562 *monitoring*, in *Proceedings of the 18th International Conference on*
563 *Automation & Computing*2012: Loughborough University, Leicestershire, UK.
- 564 22. Cross, E.J., K. Worden, and Q. Chen, *Cointegration: a novel approach for the*
565 *removal of environmental trends in structural health monitoring data*.
566 *Proceedings of the Royal Society a-Mathematical Physical and Engineering*
567 *Sciences*, 2011. **467**(2133): p. 2712-2732.
- 568 23. Cross, E.J., et al., *Features for damage detection with insensitivity to*
569 *environmental and operational variations*. Proceedings of the Royal Society
570 a-Mathematical Physical and Engineering Sciences, 2012. **468**(2148): p.
571 4098-4122.
- 572 24. Elforjani, M., S. Shanbr, and E. Bechhoefer, *Detection of faulty high speed*
573 *wind turbine bearing using signal intensity estimator technique*. Wind Energy,
574 2018. **21**(1): p. 53-69.
- 575 25. Elforjani, M. and E. Bechhoefer, *Analysis of extremely modulated faulty wind*
576 *turbine data using spectral kurtosis and signal intensity estimator*. Renewable
577 Energy, 2018. **127**: p. 258-268.
- 578 26. Hills, A.F., et al., *A novel baseline model-based technique for condition*
579 *monitoring of wind turbine components*. The British Insitute of
580 Non-Destructive Testing, 2011. **53**(8): p. 434-438.
- 581 27. Sohn, H., K. Worden, and C.R. Farrar, *Statistical damage classification under*
582 *changing environmental and operational conditions*. Journal of Intelligent
583 Material Systems and Structures, 2002. **13**(9): p. 561-574.
- 584 28. Prautzsch, H., W. Boehm, and M. Paluszny, *Bézier and B-spline techniques*.

- 585 2010, Berlin ; New York: Springer.
- 586 29. Han, Q., Z. Zhang, and B. Wen, *Periodic motions of a dual-disc rotor system*
587 *with rub-impact at fixed limiter*. Proceedings of the Institution of Mechanical
588 Engineers Part C-Journal of Mechanical Engineering Science, 2008. **222**(10): p.
589 1935-1946.
- 590 30. Billings, S.A., S. Chen, and M.J. Korenberg, *Identification of MIMO Non-Linear*
591 *Systems Using a Forward-Regression Orthogonal Estimator*. International
592 Journal of Control, 1989. **49**(6): p. 2157-2189.
- 593 31. Billings, S.A. and H.L. Wei, *An adaptive orthogonal search algorithm for model*
594 *subset selection and non-linear system identification*. International Journal of
595 Control, 2008. **81**(5): p. 714-724.
- 596 32. Marquez, F.P.G., et al., *Condition monitoring of wind turbines: Techniques and*
597 *methods*. Renewable Energy, 2012. **46**: p. 169-178.
- 598
- 599

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.