A methodology using the spectral coherence and healthy historical data to perform gearbox fault diagnosis under varying operating conditions

Stephan Schmidt[a,*], P. Stephan Heyns[a], Konstantinos C. Gryllias[b,c]

[a] *Centre for Asset Integrity Management, Department of Mechanical and Aeronautical Engineering, University of Pretoria, Pretoria, South Africa*
[b] *Department of Mechanical Engineering, PMA Division, KU Leuven, Celestijnenlaan 300, 3001 Heverlee, Belgium*
[c] *Dynamics of Mechanical and Mechatronic Systems, Flanders Make, Belgium*

Abstract

Condition monitoring is usually performed over long periods of time when critical rotating machines such as wind turbine gearboxes are monitored. There are many potential signal processing and analysis techniques that can be utilised to diagnose the machine from the condition monitoring data, however, they seldom incorporate the available healthy historical data of a machine systematically in the fault diagnosis process. Hence, a methodology is proposed in this article which supplements the order-frequency spectral coherence with historical data from a healthy machine to perform automatic fault detection, automatic fault localisation and fault trending. This has the benefit that the order-frequency spectral coherence, a very powerful technique for rotating machine fault diagnosis under varying speed conditions, can be utilised without requiring an expert to interpret the results. In this methodology, an extended version of the improved envelope spectrum is utilised to extract features from the order-frequency spectral coherence, whereafter a probabilistic model is carefully used to calculate a diagnostic metric for automatic fault detection and localisation. The methodology is investigated on numerical gearbox data as well as experimental gearbox data, both acquired under time-varying operating conditions with two probabilistic models, namely a Gaussian model and a kernel density estimator, compared as well. The results indicate the potential of this methodology for performing gearbox fault diagnosis under varying operating conditions.

*Corresponding author.
Email address: stephan.schmidt@up.ac.za (Stephan Schmidt)

Keywords:

Gearbox diagnostics, Novelty detection, Order-frequency spectral coherence,
Time-varying operating conditions

1. Introduction

Gearboxes found in wind turbines for example are subjected to harsh operating conditions, which result in the deterioration and potential failure of its bearings and gears [1, 2]. This could ultimately lead to the complete failure of the gearbox and therefore detecting and localising incipient damage are very important. It is possible to monitor the condition of the gearbox, by collecting and interrogating condition monitoring data at regular intervals [1, 3]. It is also possible to use oil and lubrication analysis, acoustic emission analysis and vibration analysis to infer the condition of the machine [1, 3]. The benefits of using vibration and acoustic emission measurements however are that they are sensitive to instantaneous changes in the condition of the machine and rich with diagnostic information. Since, gearboxes often operate under time-varying conditions which manifest in the condition monitoring data and impede the condition monitoring process, it is important to develop fault diagnosis techniques which can be used under time-varying operating conditions.

Rotating machine vibration signals are inherently cyclostationary [4], which makes cyclostationary analysis techniques well-suited for analysing the vibration signals. The squared envelope spectrum, the spectral correlation and its power normalised form, i.e. the spectral coherence, are very powerful second-order cyclostationary techniques used to perform bearing and gear fault diagnosis [5–9]. However, the vibration signals become cyclo-non-stationary under varying speed conditions, which impede the performance of cyclostationary techniques. Therefore, Abboud et al. [10, 11] extended the spectral correlation and spectral coherence for cyclo-non-stationary signals, which resulted in the order-frequency spectral correlation and the order-frequency spectral coherence to be developed for cyclo-non-stationary signals. The order-frequency spectral coherence has been very successful for diagnosing rotating machines under varying operating conditions. Abboud et al. [9] also compared different deterministic-random separation tasks for using the Squared Envelope Spectrum (SES), one of the most popular techniques for bearing

diagnostics, under time-varying operating conditions.

Condition monitoring data are usually acquired over long periods of time from a healthy gearbox and then ultimately damage develops in a machine component or components which are then replaced based on an appropriate maintenance strategy. This means that historical data of a machine in a healthy condition can be easily acquired. Combining this plentiful resource of historical data from healthy machines (referred to as healthy historical data) and the spectral coherence for diagnosing the machine under time-varying operating conditions could potentially enhance the performance of the spectral coherence even further. The potential benefit of combining historical data and the spectral coherence was illustrated in the paper by Schmidt et al. [12] on numerical gearbox data. In the aforementioned paper, a methodology is proposed using the concept of feature windows to extract features from the spectral coherence, whereafter a multivariate Gaussian model of the healthy historical features is used to perform fault detection and localisation. However, the features were not properly motivated and the Gaussian model has many parameters that need to be estimated, which ultimately restrict the number of features that can be extracted from a given dataset. Lastly, the fault detection and localisation were also performed manually in this paper.

Hence, in this article, a methodology is proposed to automatically detect and localise anomalous behaviour in specific machine components under time-varying operating conditions to address the shortcomings of the method in Ref. [12]. This is performed by firstly extracting features from the order-frequency spectral coherence with the modified Improved Envelope Spectrum (IES), whereafter a model of the historical healthy features produces a diagnostic metric that can be used to automatically detect and localise the damaged component in the machine. In summary, the contributions of this paper are as follows:

- A methodology is proposed to automatically perform fault detection and fault localisation for gearboxes operating under time-varying conditions. This can enhance the performance of the order-frequency spectral coherence even more for fault diagnosis.
- The IES, a very powerful representation for fault diagnosis, is modified to extract features from the order-frequency spectral coherence. This means that well-motivated

features can be used without performing feature engineering.

- Two models, the Gaussian model and a model using kernel density estimators (a flexible non-parametric model [13]), are compared to investigate the benefits of using a non-Gaussian model as opposed to a Gaussian model on non-Gaussian datasets.
- The methodology is validated on numerical gearbox data as well as on experimental gearbox data, both generated under time-varying operating conditions.

The layout of the paper is as follows: In Section 2, the methodology is presented in detail, whereafter the methodology is applied on numerical data in Section 3 and on experimental data in Section 4. Conclusions are drawn and recommendations are made in Section 5. Appendix A contains additional information related to the numerical gearbox model presented in Section 3.1.

2. Methodology

In this methodology, it is assumed that a historical dataset of vibration signals, denoted $\{x(t)\}$ and their corresponding rotational speed signals $\{\dot{\theta}(t)\}$ in rad/s are available from the rotating machine in a healthy condition as well as the analytical cyclic orders of the components-of-interest, e.g. the Ball-Pass Outer race Order (BPOO) and Ball-Pass Inner race Order (BPIO) of the bearing-of-interest. This methodology has two phases, namely, a training and application phase, with both presented in Figure 1. In the training phase, the

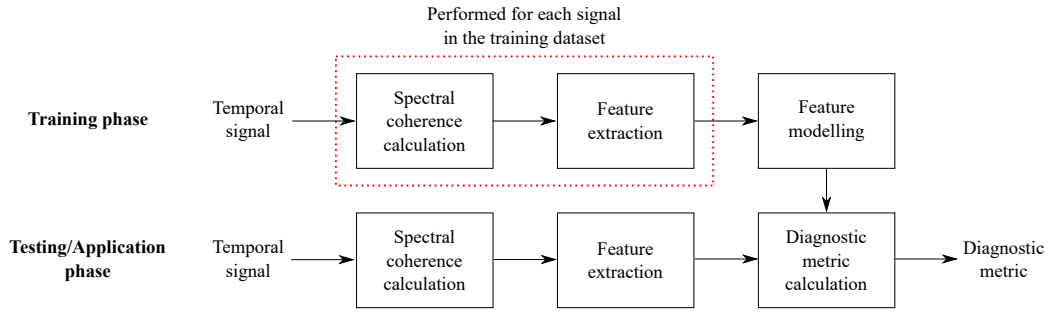

Figure 1: Process diagram of the proposed methodology.

spectral coherence of each vibration signal in the historical dataset $\{x(t)\}$ is calculated. Features are subsequently extracted from the spectral coherences of the healthy machine vibration data and modelled with a probabilistic model. In the testing or application

phase of the methodology, the probabilistic model of the healthy features is used with features extracted from the new vibration signal, to calculate a diagnostic metric which is subsequently used to infer the condition of the machine.

A more detailed overview of each step in the methodology is presented in subsequent sections.

2.1. Spectral coherence calculation

Damaged rotating machine components result in periodical impacts that generate vibration signals $x(t)$ with periodical statistics in the angle domain θ, while the impacts excite time-invariant structural resonances [11, 14]. The implication of this is that the angle-time autocorrelation function of the vibration signal [11]

$$\mathcal{R}_{2x}\left(\theta,\tau\right)=\mathbb{E}\left\{x\left(t(\theta)\right)x(t\left(\theta\right)-\tau)^{*}\right\}, \tag{1}$$

which is a function of both angle θ and time-lag τ, is periodic in the angle domain, i.e. $\mathcal{R}_{2x}\left(\theta,\tau\right)=\mathcal{R}_{2x}\left(\theta+\Theta_{c},\tau\right)$ for an angle Θ_c in radians. The cyclic order α of the impacts and the specific spectral frequencies f being excited by the impacts can be determined from the order-frequency spectral correlation. The order-frequency spectral correlation is defined as the double Fourier transform of the angle-time autocorrelation function [10, 11]

$$\mathcal{S}_{2x}\left(\alpha,f\right)=\underset{\theta\rightarrow\alpha}{\mathcal{F}}\underset{\tau\rightarrow f}{\mathcal{F}}\left(\mathcal{R}_{2x}\left(\theta,\tau\right)\right), \tag{2}$$

where $\underset{a\rightarrow b}{\mathcal{F}}$ denotes the Fourier transform from the a to the b domain, and the cyclic orders and the spectral frequency are denoted by α and f respectively. The order-frequency spectral correlation can also equivalently be written as [10]

$$\mathcal{S}_{2x}\left(\alpha,f\right)=\lim_{T\rightarrow\infty}\frac{1}{\theta\left(T\right)-\theta\left(0\right)}\mathbb{E}\left\{\mathcal{F}_{T}\left(x(t)\right)^{*}\cdot\mathcal{F}_{T}\left(x(t)e^{-j\alpha\theta(t)}\dot{\theta}(t)\right)\right\}, \tag{3}$$

where T is the time-period under consideration and $\dot{\theta}(t)=\frac{d}{dt}\theta(t)$. The spectral correlation is dependent on the power spectrum of the vibration signal and can potentially mask weak cyclostationary components in the signal and therefore the order-frequency spectral coherence [10]

$$\gamma_{2x}(\alpha,f)=\frac{\mathcal{S}_{2x}(\alpha,f)}{|\mathcal{S}_{2x}(0,f)\mathcal{S}_{2x_{\alpha}}(0,f)|^{\frac{1}{2}}}, \tag{4}$$

is used here instead of the order-frequency spectral correlation.

The Order-Frequency Spectral Correlation (OFSC) and the Order-Frequency Spectral Coherence (OFSCoh) cannot be calculated by using Equations (3) and (4) in practice, but estimators of the OFSC should be used instead. The Averaged Cyclic Periodogram (ACP) estimator of the OFSC is used in this paper due to its relatively good bias and variance properties and relatively good computational efficiency [6, 10]. Essentially, the ACP removes the expectation operator $\mathbb{E}$ in Equation (3), by performing this calculation for different segments of the vibration signal, whereafter a weighted average over the results of the different segments is calculated [10]. A detailed overview of using the ACP for estimating the spectral correlation and related representations is given in Refs. [6, 10]. It is important to note that the recent developments in the Fast Spectral Correlation (FSC) [15] and the faster estimator of the FSC presented in Ref. [16] make the potential application of this methodology more computationally efficient.

2.2. Feature extraction

We desire to use the spectral coherence with historical data to automatically infer the condition of the machine. It is possible to learn the whole spectral coherence from the healthy historical data to detect anomalous behaviour with a flexible model such as a convolutional neural network. However, it can be difficult to select and to motivate the appropriate architecture [17]. It is computationally expensive to optimise and it does not naturally provide information on which component causes the anomalous behaviour. Hence, a feature extraction procedure is used in this paper that makes it possible to not only automatically detect damage, but also to determine which component is damaged with simple probabilistic models.

We assume that there is a set of components being monitored that has a corresponding set of analytical or predetermined cyclic orders $\{\hat{\alpha}_c\}$. Hence, it is possible to extract features from the spectral coherence at the set of cyclic orders $\{\hat{\alpha}_c\}$ and their harmonics. We desire to extract features based on well developed theory and not to perform feature engineering (i.e. to optimise the features to achieve an optimal outcome). Therefore, the Improved Envelope Spectrum (IES) is used as a starting point for the feature extraction procedure. The IES [18]

$$I_{2x}(\alpha; f_1, f_2) = \frac{1}{f_2 - f_1} \int_{f_1}^{f_2} |\gamma_{2x}(\alpha, f)|^2 df, \tag{5}$$

is the average squared magnitude spectral coherence over a predefined frequency band $f \in [f_1, f_2]$ and is better suited for damage detection than the Squared Envelope Spectrum (SES) if the frequency band is selected correctly [18]. A detailed overview of the SES and its relationship to cyclostationary analysis can be found in Refs. [8, 9, 19]. It is therefore possible to define the features for a specific monitored component as follows

$$\boldsymbol{b}_c = [I_{2x}(\hat{\alpha}_c \cdot 1; f_1, f_2), I_{2x}(\hat{\alpha}_c \cdot 2; f_1, f_2), \ldots, I_{2x}(\hat{\alpha}_c \cdot N_{harm}; f_1, f_2)]^T \tag{6}$$

for the predefined frequency band $[f_1, f_2]$ and N_{harm} harmonics of the associated estimated or analytical cyclic order $\hat{\alpha}_c$. However, it is not possible to determine the optimal frequency band $[f_1, f_2]$ from only healthy historical data (i.e. to estimate it during the training phase of the methodology) and the available analytical estimate of α_c, denoted $\hat{\alpha}_c$, could be slightly different from the actual cyclic order and/or the analytical estimate may not be defined for the estimated OFSCoh (i.e. the estimated OFSCoh is defined for a discrete set of cyclic orders α and spectral frequencies f). Hence, due to the uncertainty over the actual α_c, it is best to describe it with a probability density function $p_{\alpha_c}(\alpha_c|\hat{\alpha}_c)$. The subscript α_c in $p_{\alpha_c}(\alpha_c|\hat{\alpha}_c)$ is used to avoid confusion between this model and the models used in Section 2.3.

The expected value of the bivariate improved envelope spectrum

$$\mathbb{E}_{\alpha_c}\left\{I_{2x}(\alpha_c, f_m; \Delta f)|\, \hat{\alpha_c}\right\} = \mathbb{E}_{\alpha_c}\left\{ \frac{1}{\Delta f}\int_{f_m-\Delta f/2}^{f_m+\Delta f/2} |\gamma_{2x}(\alpha_c, f)|^2 df \,\middle|\, \hat{\alpha_c} \right\} \tag{7}$$

$$= \int_{\alpha_c} \frac{1}{\Delta f}\int_{f_m-\Delta f/2}^{f_m+\Delta f/2} |\gamma_{2x}(\alpha_c, f)|^2 df \cdot p_{\alpha_c}(\alpha_c|\hat{\alpha}_c) d\alpha_c. \tag{8}$$

is calculated to overcome the need to determine the unknown frequency band $[f_1, f_2]$ a priori, because it is now parametrised in terms of cyclic orders and spectral frequencies. The uncertainty in the cyclic order is also incorporated in Equation (7) by calculating the expected value of the IES over the random cyclic order variable, while being conditioned on the analytical estimate $\hat{\alpha}_c$.

For simplicity, it is assumed that $p_{\alpha_c}(\alpha_c|\hat{\alpha}_c)$ is a uniform distribution with a domain of $\alpha_c \in [\hat{\alpha}_c - \Delta\alpha/2, \hat{\alpha}_c + \Delta\alpha/2]$ and therefore Equation (8) can be simplified to

$$J_{2x}(\hat{\alpha}_c, f_m; \Delta f, \Delta\alpha) = \frac{1}{\Delta f \cdot \Delta\alpha}\int_{\hat{\alpha}_c-\Delta\alpha/2}^{\hat{\alpha}_c+\Delta\alpha/2}\int_{f_m-\Delta f/2}^{f_m+\Delta f/2} |\gamma_{2x}(\alpha, f)|^2 df \; d\alpha, \tag{9}$$

where the notation J is used to avoid confusion over the I used in Equation (5). The cyclic order tolerance $\Delta\alpha$ and the integration points, i.e. f_i and Δf, needs to be specified before applying this methodology. Therefore, as opposed to using Equation (6), the following features, $\boldsymbol{b}$, are used for characteristic c

$$\boldsymbol{b}_c = \left[J_{2x}(\hat{\alpha}_c \cdot 1, f_1), \ldots, J_{2x}(\hat{\alpha}_c \cdot k, f_m), \ldots, J_{2x}(\hat{\alpha}_c \cdot N_{harm}, f_{N_f})\right]^T, \tag{10}$$

and is defined for all combinations of $k = 1, 2, \ldots N_{harm}$ and $m = 1, 2, \ldots, N_f$, where N_{harm} and N_f denote the number of harmonics and number of frequency bands being used respectively, i.e. $\boldsymbol{b}_c \in \mathbb{R}^{(N_f \cdot N_{harm}) \times 1}$. Equation (9) is equivalent to using a feature window on the squared-magnitude OFSCoh, whereafter the mean of the windowed data is used as a feature [12].

Ultimately, the features of all characteristics are concatenated together to form the N_{feat} features of a specific measurement

$$\boldsymbol{b} = [\boldsymbol{b}_1^T, \cdots, \boldsymbol{b}_c^T, \ldots \boldsymbol{b}_{N_c}^T]^T, \tag{11}$$

where N_c denotes the number of components being monitored and $\boldsymbol{b} \in \mathbb{R}^{N_{feat}}$.

2.3. Feature modelling

We aim to model the features of a healthy machine with a probabilistic model $p(\boldsymbol{b})$, which can ultimately be used to detect damage. The distribution of the healthy features of the gearbox is unknown and therefore two different modelling approaches, namely, a Gaussian model and a kernel density estimator, are compared.

2.3.1. Gaussian model

Even though the features are not expected to be Gaussian, it is naively assumed that the features are Gaussian distributed and therefore a Gaussian distribution is used to model the data. The multivariate Gaussian model, described by a mean and a covariance parameter, can be used to detect anomalous behaviour in the machine [12, 20]. However, due to the large number of features that can potentially be encountered and large number of parameters that need to be estimated for the mean and the covariance, it is rather assumed that the features are identically and independently distributed (i.i.d) if the gearbox is

healthy, which results in the following simplified model

$$p(\boldsymbol{b}; \boldsymbol{\mu}_b, \boldsymbol{\sigma}_b) = \prod_{n=1}^{N_{feat}} \mathcal{N}(b_n; \mu_{b_n}, \sigma_{b_n}^2), \tag{12}$$

where b_n denotes the nth feature and μ_{b_n} and $\sigma_{b_n}^2$ denote the corresponding mean and variance of the univariate Gaussian model $\mathcal{N}$. The same assumption is made by a Gaussian Naive-Bayes classification model, which has performed very well despite these seemingly limiting assumptions [13]. In Equation (12), the empirical mean and variance of the nth feature of the healthy data are used as the mean μ_{b_n} and the variance $\sigma_{b_n}^2$ of the Gaussian model.

However, as the Gaussian model is unable to capture the underlying distribution of non-Gaussian data, it could potentially be insensitive to incipient damage and therefore a non-parametric density estimation method is considered in the next section.

2.3.2. Kernel density model

Kernel density estimators are non-parametric models of the data, with the benefit that they do make fewer assumptions on the shape of the distribution and therefore can potentially describe the actual distribution better than the Gaussian model [13]. However, this flexibility comes with the risk of overfitting on the training data and therefore care should be taken to select the parameters. The kernel density estimator of the assumed independently and identically distributed healthy features [13]

$$p(\boldsymbol{b}; h) \approx \prod_{n=1}^{N_{feat}} \frac{1}{N_{train}} \sum_{i=1}^{N_{train}} \frac{1}{h} \cdot \mathcal{K}\left(\frac{b_n - b_n^{(i)}}{h}\right), \tag{13}$$

is calculated for a specific kernel function $\mathcal{K}(\cdot)$ and hyperparameter h. The widely used Gaussian kernel, which enforces smoothness in the estimation [13], is used and results in the following probability density function:

$$p(\boldsymbol{b}; h) \approx \prod_{n=1}^{N_{feat}} \frac{1}{N_{train}} \sum_{i=1}^{N_{train}} \frac{1}{\sqrt{2\pi}h} \cdot \exp\left(-\frac{1}{2}\left(\frac{b_n - b_n^{(i)}}{h}\right)^2\right). \tag{14}$$

In Equation (14), the training feature set contains N_{train} samples from the N_{feat}-dimensional features of a healthy gearbox, with $b_n^{(i)}$ denoting the nth feature of the ith measurement in the training dataset.

The hyperparameter h in Equation (14) cannot be selected by maximising the likelihood for the training dataset and therefore a k-fold cross-validation process is used. In a k-fold cross validation procedure, the training data are segmented into k equally-sized datasets, whereafter $k-1$ datasets are used to train the model and the remaining dataset is used as a validation dataset to calculate a score or prediction error of the model [13]. This process is repeated for all possible combinations of training and validation data, with the h that maximises the likelihood of the validation datasets being used in Equation (14) for subsequent predictions. Bishop [21] used a kernel density estimator for novelty detection, where the parameter h was selected as the average distance between each point in the dataset and its ten nearest neighbours.

In the next section, the development of diagnostic metrics for automatic fault detection is discussed. For notational simplicity, the dependence of the probability density functions on the hyperparameters is omitted in subsequent sections.

2.4. Diagnostic metric calculation and automatic fault detection

The model of the healthy dataset can be used to develop a diagnostic metric, i.e. a metric which can be used to infer the condition of the machine. It is expected that the likelihood of the parameters of a healthy model would decrease as the machine transitions from a healthy condition to a severely damage condition. Hence, the negative logarithm of the model $-\log p(\boldsymbol{b})$ is appropriate as a diagnostic metric.

However, $-\log p(\boldsymbol{b})$ can only be used to detect changes in the system and not to detect changes in a specific component of the machine. Therefore, the negative logarithm of the marginal distribution over characteristic c, i.e. $-\log p(\boldsymbol{b}_c)$, is also used here as a diagnostic metric. This allows one to not only detect changes in the condition of the machine, but also to detect changes in the condition of a specific component. The marginal distribution can either be obtained by integrating the distribution $p(\boldsymbol{b})$ over the complementary features of $\boldsymbol{b}_c$, or by directly modelling $p(\boldsymbol{b}_c)$. It is easy to calculate the marginal distribution from the full distribution when the i.i.d. assumption is made or when a Gaussian model is used [13]. Both diagnostic metrics, i.e. $-\log p(\boldsymbol{b})$ and $-\log p(\boldsymbol{b}_c)$, are evaluated in subsequent investigations.

It is subsequently possible to calculate an alarm threshold β_{thres} from the diagnostic

metrics of the healthy data, whereafter the threshold can be used to determine whether anomalous behaviour is present or not. For example, $-\log p(\boldsymbol{b}_c) > \beta_{thres}$ can be used for detecting anomalous behaviour or $-\log p(\boldsymbol{b}_c) \leq \beta_{thres}$ for detecting normal behaviour. In this paper, the 99th percentile of the diagnostic metric of the healthy data is used as a threshold β_{thres} for detection. The motivation for using the 99th percentile as a threshold is as follows: It is desired to be able to detect the faults early in the degradation process to ensure that there is sufficient time to properly trend the components and to perform maintenance planning (e.g. spare parts procurement). The 99th percentile is calculated-based on the empirical distribution of the data and not based on a naive assumption, e.g. $\mu+3\sigma$, and would therefore be very sensitive to subtle changes in the data. Unfortunately, if this threshold is compared to the raw data, many false alarms would be triggered as well and therefore the threshold is rather compared to the median of the diagnostic metrics of the previous N measurements in this methodology. The median is robust to outliers and therefore if the threshold is exceeded, an anomaly would be present. If this methodology was applied on in industrial application, it would be sensible to define different thresholds, where the 99th percentile is only used to warn the maintenance team and other less conservative thresholds being used to stop the machine.

2.5. Final remarks about the methodology

In summary, the following information needs to be provided or specified when implementing this methodology:

- A dataset of vibration signals and their corresponding rotational speed signals from a healthy machine.
- The analytical cyclic orders of the components-of-interest $\{\hat{\alpha}_c\}$ as well as the number of harmonics N_{harm} that need to be monitored.
- A cyclic order tolerance $\Delta\alpha$ of the cyclic orders of the components-of-interest.
- The bounds of the integration Δf and the spectral integration points f_i used in the calculation of the features.

The methodology is investigated on numerical gearbox data in Section 3 and on experimental gearbox data in Section 4.

3. Numerical investigation

A phenomenological gearbox model, presented in the paper by Abboud et al. [9] and also used by Schmidt et al. [20] is used in this section, because it is capable of modelling complex phenomena such as the amplitude variation in vibration signals due to variations in rotational speed, while being computationally efficient to calculate.

3.1. Numerical model

The casing vibration signal of the numerical gearbox model

$$x(t) = x_{gmc}(t) + x_b(t) + x_{dgd}(t) + x_n(t), \tag{15}$$

is decomposed in terms of a gear mesh component $x_{gmc}(t)$, an outer race bearing damage component $x_b(t)$, a distributed gear damage component $x_{dgd}(t)$ and a noise component $x_n(t)$. The bearing and gear damage are explicitly introduced in the casing vibration signal by decomposing the associated signal component i into a fault severity factor FS_i and a baseline signal component $\tilde{x}_i(t)$

$$x(t) = x_{gmc}(t) + FS_b \cdot \tilde{x}_b(t) + FS_{dgd} \cdot \tilde{x}_{dgd}(t) + x_n(t). \tag{16}$$

All combinations of $\{FS_b\} = \{0, 1.33, 2.66, 4\}$ and $\{FS_{dgd}\} = \{0, 1.33, 2.66, 4\}$ are investigated in this work, where $FS_b = 0$ and $FS_{dgd} = 0$ define a healthy gearbox. The gear mesh component in the casing vibration signal

$$x_{gmc}(t) = h_{gmc}(t) \otimes \left|\dot{\theta}(t)\right|^2 \cdot \sum_{k=1}^{N_{gmc}} A_n^{gmc} \sin\left(k \cdot \theta(t) \cdot N_1 + \phi_n^{gmc}\right) \tag{17}$$

is obtained by filtering the signal at the source with the impulse response function $h_{gmc}(t)$. In Equation (17), A_k^{gmc} and ϕ_k^{gmc} denote the amplitude and the phase of the kth harmonic of the gear mesh component, and $\left|\dot{\theta}(t)\right|^2$ simulates the amplitude variation of the signal due to speed variations. The distributed gear damage component

$$\tilde{x}_{dgd}(t) = h_{dgd}(t) \otimes \left|\dot{\theta}(t)\right|^2 \cdot \epsilon(t) \cdot \sum_{k=1}^{N_{gmc}} A_n^{dgd} \cdot \sin\left(k \cdot \theta(t) + \phi_n^{dgd}\right), \tag{18}$$

has a similar form to Equation (17), except that it is multiplied by an additional zero mean Gaussian random variable $\epsilon(t)$ and its fundamental frequency is 1.0 shaft orders.

The outer bearing component consists of a train of impulses

$$\tilde{x}_b(t) = h_b(t) \otimes \left|\dot{\theta}(t)\right|^2 \sum_{n=1}^{N_b} A_n^b \cdot \delta\left(t - \mathcal{T}_n\left(\theta(t)\right)\right) \tag{19}$$

multiplied by a uniformly distributed random component A_n^b and a component which scales the amplitude of the bearing impulse with the rotational speed. The time-of-arrival of the nth impulse $\mathcal{T}_n$ is dependent on the instantaneous phase of the system $\theta(t)$.

Lastly, the broadband noise component

$$x_n(t) = \left|\dot{\theta}(t)\right|^2 \cdot \epsilon(t), \tag{20}$$

contains a zero mean Gaussian component $\epsilon(t)$, which is scaled with the rotational speed of the system. In all cases, a single degree-of-freedom impulse response function is used for $h_i(t)$ and is described by two parameters, the natural frequency and the damping ratio [20]. The data are generated with a sampling frequency of 20 kHz.

The fundamental order of the distributed gear damage is 1.0 and the natural frequency and the damping ratio of its impulse response function are 3000 Hz and 0.05, respectively. The fundamental order of the outer race bearing component is 2.57 and the natural frequency and the damping ratio of its impulse response function are 7000 Hz and 0.05, respectively.

The operating conditions that can for example be encountered in the gearboxes of wind turbines and bucket wheel excavators are inherently non-stationary [22–24]. We therefore desire to simulate data under time-varying operating conditions where the rotational speed of each measurement is different. Instead of defining many unique rotational speed signals, three operating condition profile functions were used as baseline rotational speed signals, whereafter the phase of the signals and the amplitudes of the components were sampled from probability density distributions. This ensured that each casing vibration signal is unique and that the training dataset does not comprise of exactly the same rotational speed signals as the testing dataset. The resulting rotational speeds are presented in Figure 2 for the three rotational speed profiles. A detailed description of the procedure used to generate the signals in Figure 2 is given in Appendix A.

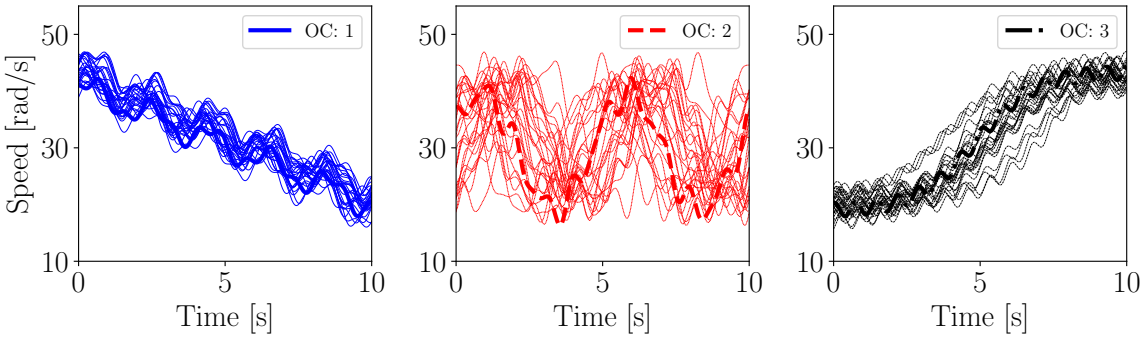

Figure 2: The Operating Conditions (OCs) that are used in the simulated model. Three profiles are used to generate the samples presented in the figure, with a sample of each profile highlighted with a thicker line and all samples associated with a specific profile having the same colour.

3.2. Results

The methodology was implemented by monitoring the gear for distributed gear damage and the bearing for outer race damage. Due to the discrete nature of the estimated OFSCoh, the monitored cyclic orders are $\{\alpha_c\} = \{0.975,\ 2.550\}$ with 3 harmonics used, i.e. $N_{harm} = 3$. The spectral frequency resolution in Equation (9), denoted Δf, is selected to ensure that it allows changes in specific frequency bands to be detected, while allowing Equation (9) to be properly estimated. The $\Delta\alpha$ is selected to include at least three samples of the OFSCoh along the α-axis. The resulting parameters are $\Delta f = 781.25$ Hz and $\Delta\alpha = 0.1$ shaft orders.

The OFSCoh for a case with bearing and distributed gear damage is presented in Figure 3(i) with the centres of the integration areas $(k \cdot \alpha_c, f_m)$ and their integration bound $[\alpha_c - \Delta\alpha/2, \alpha_c + \Delta\alpha/2]$ in Equation (9) are presented in Figure 3(ii). As a result, 72 features are extracted from the spectral coherence of each dataset, i.e. $N_{feat} = 72$. The results of the two different modelling approaches, discussed in Section 2.3, are presented separately in Sections 3.2.1 and 3.2.2. The healthy dataset in this section comprises of 12 measurements of each operating condition as presented in Figure 2, i.e. the training dataset comprises of 36 measurements with $FS_b = FS_{dgd} = 0$ in Equation (16).

3.2.1. Gaussian model

The mean and the variance of the Gaussian model for a specific feature are set to the empirical mean and variance of the corresponding feature extracted from the healthy

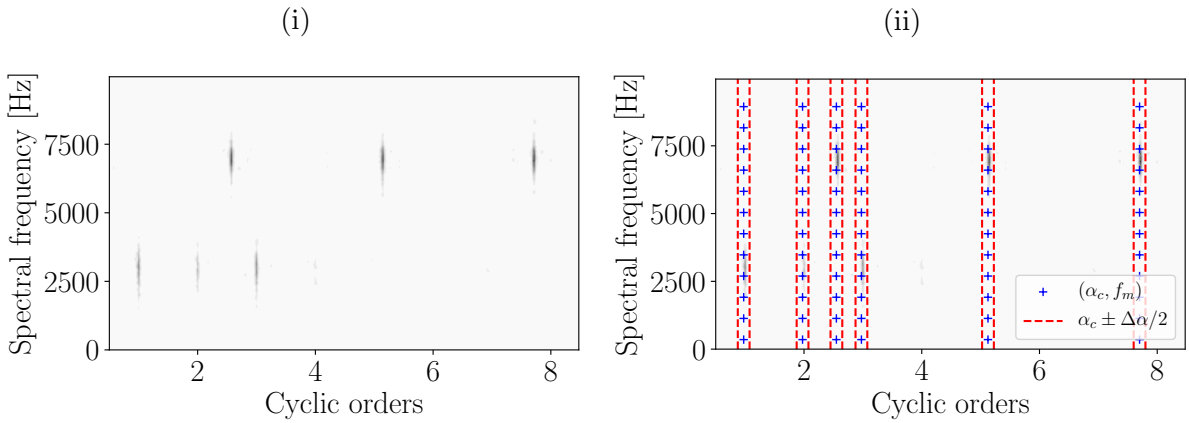

Figure 3: The spectral coherence of the casing vibration signal from the numerical gearbox are presented in Figure 3(i) for $FS_b = 4$ and $FS_{dgd} = 4$. The integration region is presented in Figure 3(ii). In Figure 3(ii), the Δf bound is not shown, because Δf is fixed and there is no overlap and gaps between vertically coinciding integration bounds.

gearbox data. The diagnostic metric for the complete feature set of a specific measurement $-\log p(\boldsymbol{b})$ is calculated with Equation (12) for the different combinations of FS_i and is presented as a three-dimensional scatter plot in Figure 4(i). A contour, which indicates the relative change of the magnitudes of the diagnostic metric as the condition changes from a healthy system, is presented as well. The diagnostic metric $-\log p(\boldsymbol{b})$ changes significantly with changes in FS_b, while changing less for changes in FS_{dgd}.

A threshold β_{thres} was calculated from the 99th percentile of the diagnostic metric $-\log p(\boldsymbol{b})$ of the healthy data and compared to the median diagnostic metric of a specific machine condition to obtain an alarm trigger (i.e. 0 indicates that healthy behaviour is observed, while 1 indicates that anomalous behaviour is present) for all combinations of FS. It is sensible to calculate the median diagnostic metric of a few measurements of a gearbox in the same condition, to ensure that the alarm is not triggered by outliers, but by actual changes in the data. The results are presented in Figure 4(ii), where it is seen that there is only a single false negative for $FS_b = 1.33$. This is attributed to the fact that the bearing damage is very small and the Gaussian model does not perfectly describe the underlying distribution. It is expected that if a more appropriate density is used, then the results would be improved.

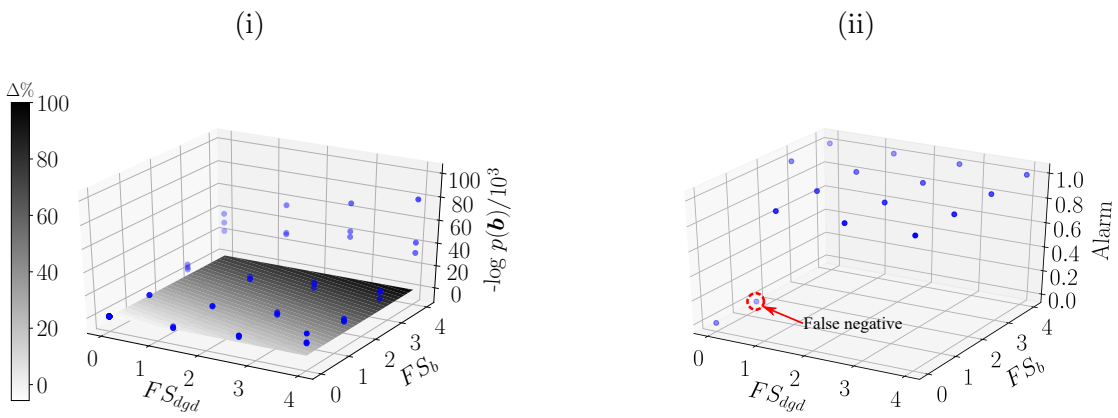

Figure 4: The diagnostic metric of the Gaussian model for the complete features $\boldsymbol{b}$, $-\log p(\boldsymbol{b})$, is presented in Figure 4(i) over the different combinations of FS_{dgd} and FS_b in a scatter plot. A contour plot of the diagnostic metric is also presented with a corresponding colorbar to indicate the trend of the diagnostic metric over the feature space. The contour is obtained by firstly fitting a two-dimensional linear regression model to the data and then to predict the relative change of the diagnostic metric from a healthy system. The alarm, after comparing the median of the data in Figure 4(i) to an alarm threshold, is presented in Figure 4(ii).

The marginal distribution of each characteristic $-\log p(\boldsymbol{b}_c)$ is calculated separately and presented in Figure 5 for the different combinations of gear and bearing damage. The presented alarm threshold is set to the 99th percentile of the diagnostic measure of the healthy dataset. It is clear from Figure 5(i) that changes in the condition of the gear result in changes in the diagnostic metric $-\log p(\boldsymbol{b}_{dgd})$, while it is largely unaffected by the presence of bearing damage. The same conclusion can be drawn from the results of the bearing diagnostic metric $-\log p(\boldsymbol{b}_b)$ in Figure 5(ii). Even though the diagnostic metrics have non-linear relationships with FS in Figure 5, the resulting change in the diagnostic metric $-\log p(\boldsymbol{b}_b)$ for $FS_b = 1.33$ for a healthy system is not prominent. This insensitivity of the Gaussian model's diagnostic metric to incipient damage is attributed to the fact that the Gaussian distribution is incapable of capturing the true underlying distribution of the features and could therefore be insensitive to small changes in the data. However, the diagnostic metric of the Gaussian model performs very well compared to the SES of the raw vibration signal. In the SES, the damaged components could not be detected for all the investigated FSs due to the low signal-to-noise ratio of the diagnostic components.

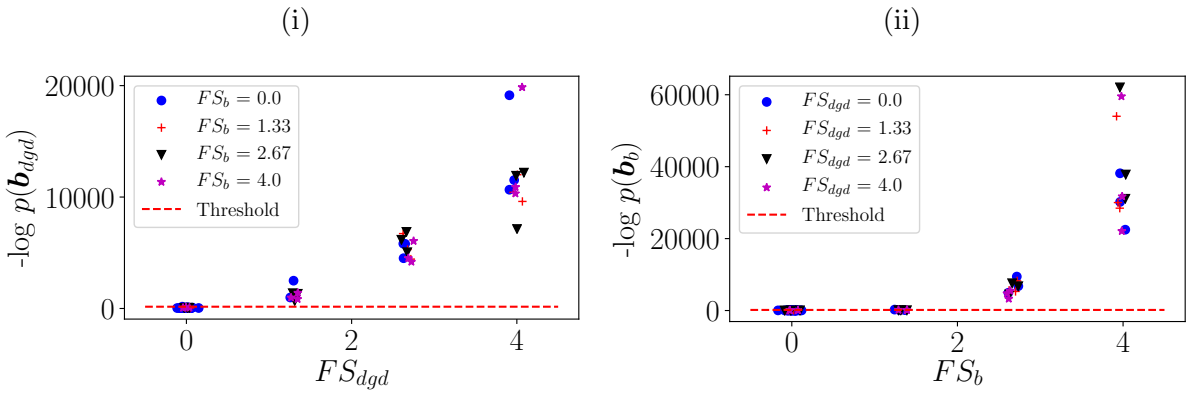

Figure 5: The diagnostic metric of the Gaussian model for the marginal distribution, i.e. associated with each monitored characteristic, is presented over the corresponding fault severity values in Figure 5(i) for the gear and in Figure 5(ii) for the bearing. An alarm threshold, calculated from the 99th percentile of the healthy data of the specific component, is also presented on the figures.

The thresholds are compared to the diagnostic data in order to obtain the alarm level in Figure 6 for automatic fault detection. The results in Figure 6 corroborate the observations

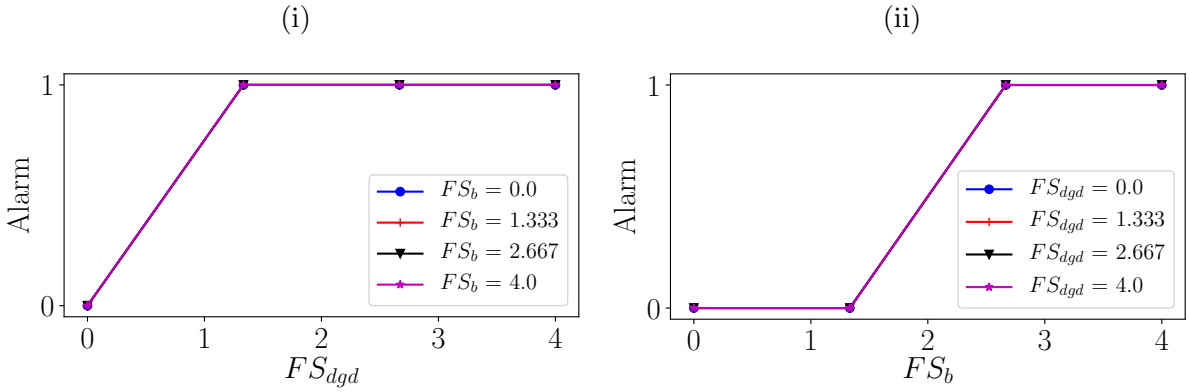

Figure 6: The alarm value is presented for the gear and bearing in Figures 6(i) and 6(ii), respectively for the data and threshold presented in Figures 5(i) and 5(ii) respectively.

made in Figure 5. The diagnostic metric of the Gaussian model is sensitive to changes in the condition of the machine, but the bearing damage at a $FS_b = 1.33$ is undetected. The results indicate that the Gaussian model performs very well despite the naive assumption that the features are Gaussian distributed.

3.2.2. Kernel density model

The results, obtained by using the kernel density model instead of the Gaussian model, are presented for the same data used in the previous section. The hyperparameter of the Gaussian kernel density estimator is obtained by defining a grid of potential h values on a logarithmic scale between $[10^{-5}, 10^{5}]$ and by selecting the appropriate h in Equation (14) which maximises the likelihood function after performing a five-fold cross validation.

The diagnostic metric $-\log p(\boldsymbol{b})$ is presented over the grid of the FS_i in Figure 7(i) with the alarm, after comparing the median of the data for a gearbox in a specific condition to an alarm threshold, presented in Figure 7(ii). The results in Figure 7(i) indicate that the

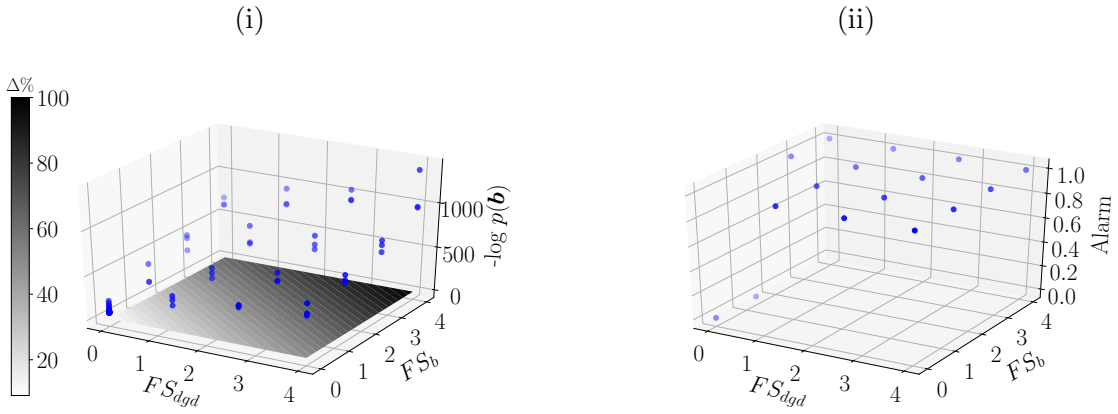

Figure 7: A scatter plot of the diagnostic metric $-\log p(\boldsymbol{b})$ for the kernel density estimator and a contour of the diagnostic metric are presented in Figure 7(i) for different combinations of fault severity factors FS_i. In Figure 7(ii) the alarm value, obtained by comparing the median data of each combination of FS_b and FS_{dgd} in Figure 7(i) to the alarm threshold, is presented similarly to Figure 4.

kernel density model is sensitive to changes in the condition of the gear and the bearing, however, its detection capabilities in Figure 7(ii) are very similar to the results in Figure 4.

Even though the diagnostic metric of the complete feature set $-\log p(\boldsymbol{b})$ performed well, the diagnostic metric from the marginal distribution contains more useful information. The marginal diagnostic metrics are presented in Figure 8(i) for the gear and in Figure 8(ii) for the bearings, with the alarm thresholds superimposed as well. It is evident from the results in Figure 8, compared to the results in Figure 5, that the kernel density model is

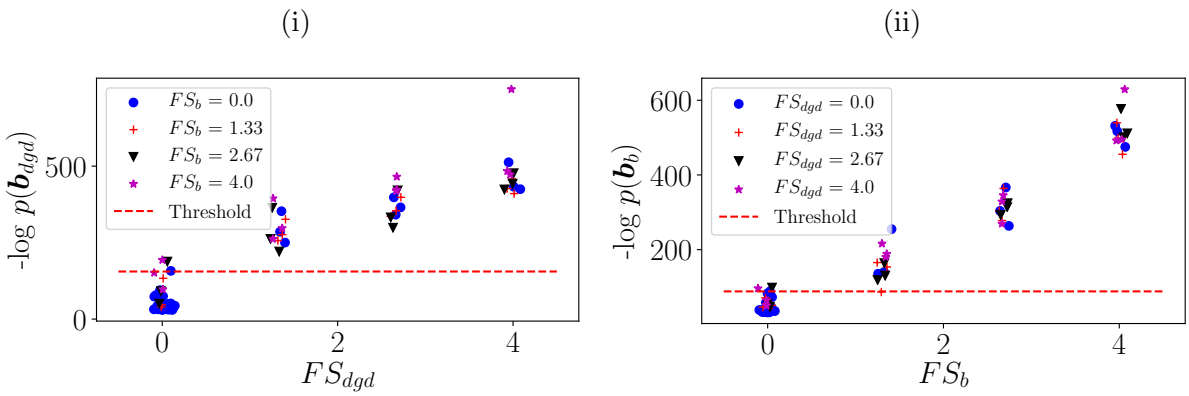

Figure 8: The diagnostic metric of the kernel density model for the marginal distribution is presented over the corresponding fault severity values in Figure 8(i) for the gear and in Figure 8(ii) for the bearing. The alarm thresholds, calculated from the healthy data of a specific diagnostic metric, are also presented in the figures.

more sensitive to incipient damage than the Gaussian model. The diagnostic metric in this case almost has a linear relationship with FS_i as opposed to the non-linear relationship of the Gaussian model, which indicates that it could be better suited for incipient damage detection. Some outliers in Figure 8(i) for $FS_{dgd} = 0$ and in Figure 8(ii) for $FS_b = 0$ exceed the threshold, which emphasise that the statistics of the data (e.g. the median of the data) and not the data itself must be compared to the threshold to avoid false alarms.

The alarm threshold in Figure 8 is compared to the median of the data from a gearbox in a specific condition, i.e. a specific combination of FS_b and FS_{dgd}, in Figure 8 to obtain the alarm value presented in Figure 9. The results in Figure 9 confirm that the kernel model is able to automatically detect the presence of the gear and the bearing damage and when compared to the results in Figure 6 it is clear that the kernel density model is more sensitive to damage than the Gaussian model. This is attributed to the fact that the kernel density model is capable of describing the actual distribution of the data more accurately than the fixed, and potentially limiting, shape of the Gaussian distribution. However, even though the Gaussian assumption is inappropriate, very good results are obtained with the Gaussian model as well.

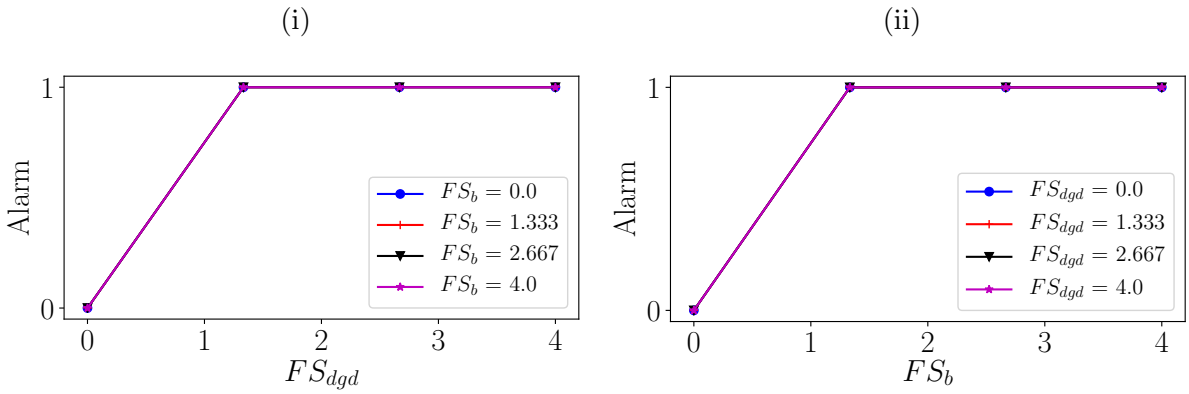

Figure 9: The alarm value is presented for the gear and bearing in Figures 9(i) and 9(ii), respectively for the data and threshold presented in Figure 8.

4. Experimental investigation

It is important to investigate the methodology on experimental data as well, therefore, the methodology is applied to gearbox data acquired under varying operating conditions. An overview of the experiments is given in Section 4.1, whereafter the results are presented in Section 4.2.

4.1. Overview of the experiments

The experimental setup is presented in Figure 10 and consists of three gearboxes, an alternator and an electrical motor. The test gearbox in Figure 10 is monitored for damage, while the alternator and the electrical motor are separately controlled to exert time-varying loads and speeds on the system. The monitored gearbox is a single stage

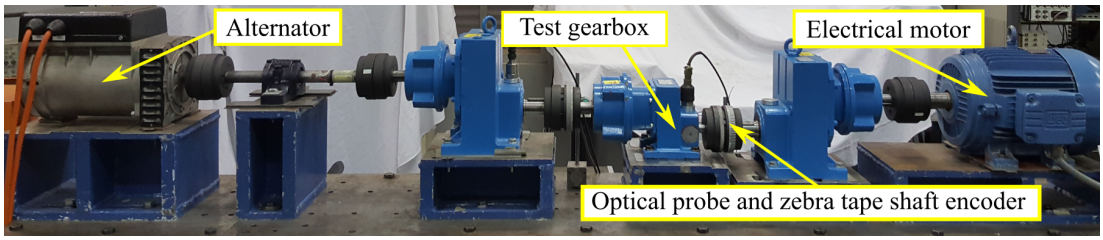

Figure 10: The experimental setup used to generate the experimental dataset.

helical gearbox with a gear ratio of 1.85 and monitored with a tri-axial accelerometer, located on the bearing housing of the input shaft of the test gearbox. The vibration signal

was acquired at a sampling frequency of 25.6 kHz with the axial component of the tri-axial accelerometer used in this investigation. The rotational speed measurements are acquired from the input shaft of the gearbox with an optical probe and a zebra tape shaft encoder. The zebra tape shaft encoder resulted in 88 pulses to be generated per revolution and its signal was acquired at a sampling frequency of 51.2 kHz. The operating conditions applied to the monitored gearbox over the 20 second measurement period are fixed for each measurement and are presented in Figure 11.

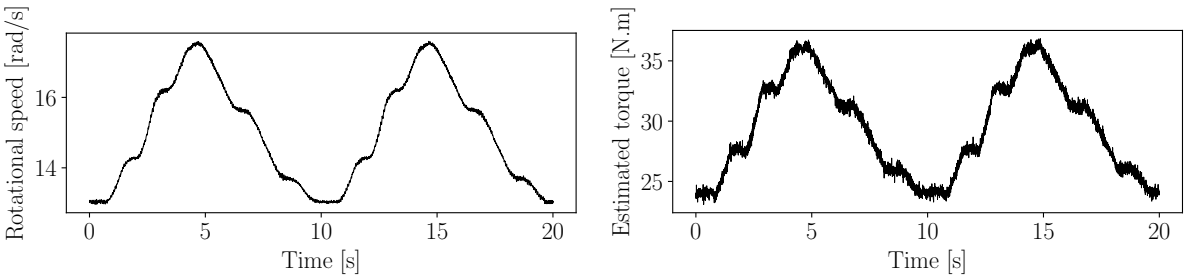

Figure 11: The operating conditions that were applied to the input shaft of the test gearbox.

Condition monitoring data were acquired from gears in a healthy condition, whereafter the gearbox was opened and the gear was damaged by seeding a slot with spark erosion in its root. The gearbox was assembled with the damaged gear, whereafter it was operated until the gear tooth failed after approximately 20 days. The damaged gear with the seeded damage is presented in Figure 12(i), while the gear with the broken tooth is presented in Figure 12(ii). Even though the gear damage in Figure 12(i) may seem large, it has little effect on the gear mesh stiffness due to the fact that helical gears are used and is therefore not easy to be detected under the varying operating conditions. In this paper, 35 measurements from the healthy machine define the healthy dataset, while 100 equally spaced measurements of the actual 1400 measurements from the damaged gearbox are used as the testing dataset.

In the next section, the results of the methodology are investigated for the Gaussian model and the kernel density model.

4.2. Results

In this investigation, the monitored gearbox components are the gear and the pinion, with the cyclic orders of the gear being 1.0 and the pinion 1.85. However, due to the

(i) Before (ii) After

Figure 12: The damaged gear with the initial damage before the test and the damaged gear with the broken tooth after the test was completed, are presented in Figure 12(i) and Figure 12(ii), respectively.

fixed grid of the OFSCoh, the corresponding set of monitored cyclic orders that can be achieved is $\{\alpha_c\} = \{1.0, 1.84\}$. The OFSCoh of the damaged gearbox is presented in Figure 13(i), while the feature windows are superimposed on the OFSCoh in Figure 13(ii). Metrics, similar to those used in Section 3.2, are used to design the integration region to

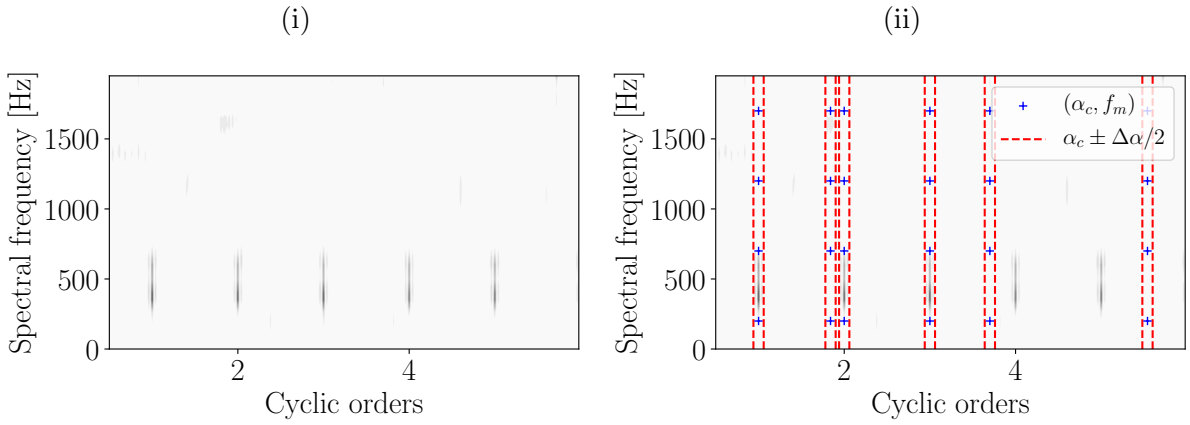

Figure 13: The OFSCoh is presented in Figure 13(i) for the damaged gear and presented in Figure 13(ii) with the integration regions used in Equation (9) superimposed on the same OFSCoh.

obtain the features. The integration bounds are given by $\Delta\alpha = 0.06$ shaft orders and $\Delta f = 500$ Hz, while three harmonics of the monitored gear and pinion components are used. This resulted in a total of 24 features being extracted from the spectral coherence

for this investigation.

In the next section, the Gaussian model is used in the methodology to model the features, whereafter the kernel density model is used in Section 4.2.2.

4.2.1. Gaussian model

The Gaussian model is implemented as in Section 3.2.1, with the raw diagnostic metric $-\log p(\boldsymbol{b})$, i.e. calculated for each measurement without processing, being presented in Figure 14(i) as a function of measurement number. It is difficult to observe changes in the

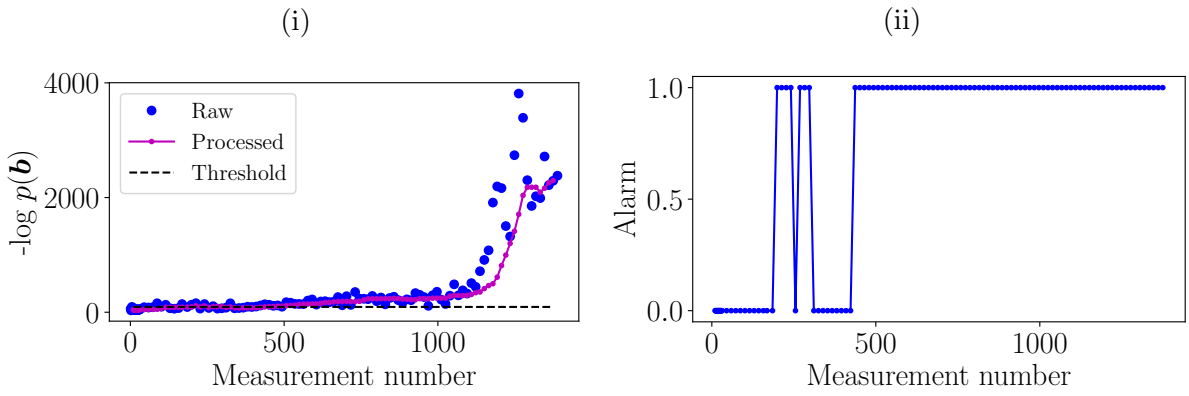

Figure 14: The diagnostic metric for the complete feature set, $-\log p(\boldsymbol{b})$, is presented for the Gaussian model for the gearbox data as a function of measurement number in Figure 14(i). The processed diagnostic metric, by calculating the median of the diagnostic metrics of consecutive measurements, is also presented in Figure 14(i) with the alarm threshold superimposed. The processed data are compared to the alarm threshold to obtain the alarm presented in Figure 14(ii).

condition of the machine for the initial measurements with large change only observed in Figure 14(i) during the final stages of the experiment.

It is necessary to compare the data to an alarm threshold for automatic fault detection. If the raw data in Figure 14(i) are compared to the alarm threshold, outliers can easily trigger false alarms. Hence, it is assumed that the condition of the gearbox remains approximately constant for two days and therefore the median of the diagnostic metrics associated with this time period is calculated and compared to an alarm threshold for fault detection. Hence, a moving median filter is applied to the raw diagnostic metrics to obtain the processed diagnostic metric, which is also presented in Figure 14(i). The processed

diagnostic metric is significantly smoother than the raw diagnostic metric and is therefore more robust to outliers. The 99th percentile of the diagnostic metric $-\log p(\boldsymbol{b})$ of the first 20 measurements is used to define the alarm threshold presented in Figure 14(i) as well.

The alarm, obtained by comparing the processed diagnostic metric to the alarm threshold, is presented in Figure 14(ii) for the data in Figure 14(i). The alarm is triggered at measurement number 199, whereafter the alarm is momentarily 0 between measurement numbers 310 and 423. Even though the alarm is correctly triggered (i.e. the gear is damaged), the subsequent alternation between triggered and non-triggered can lead to confusion.

The raw diagnostic metric from the marginal distribution is presented in Figure 15(i) for the gear and the pinion as a function of measurement number. The processed diagnostic

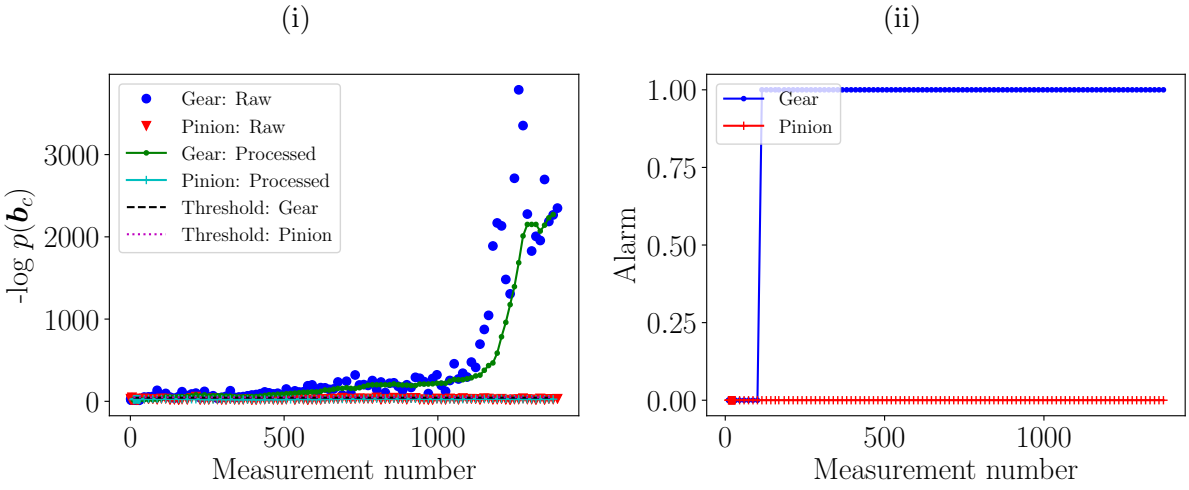

Figure 15: The raw diagnostic metric of the marginal Gaussian distribution of each monitored component $-\log p(\boldsymbol{b}_c)$ is presented in Figure 15(i). The processed diagnostic metrics, obtained with the same procedure as Figure 14(i), are presented with the alarm threshold of each monitored component as well. The alarm, by comparing each processed diagnostic metric to its corresponding alarm threshold, is presented in Figure 15(ii).

metrics and the alarm thresholds of the gear and the pinion are obtained with the same procedure as presented in Figure 14(i). The alarm thresholds of the gear and the pinion are very close to each other and their values are quite small compared to the scale of the diagnostic metric over the life of the machine. This is attributed to the little variance

in the diagnostic metrics of the initial measurements and the exponential growth of the diagnostic metric with changes in the condition of the machine.

The alarms, obtained by comparing the processed data to their respective thresholds in Figure 15(ii), are presented in Figure 15(ii). The results indicate that the pinion is in a healthy condition, while the gear damage is detected at measurement number 115. Therefore, the results are significantly better than the results in Figure 14, because the damage is detected earlier, the damaged component can be identified and the alarm is less confusing, i.e. after the alarm is triggered for the first time, it remains triggered.

4.2.2. Kernel density model

The model obtained from the kernel density estimator is applied on the same data as the Gaussian model in the previous section, with the hyperparameter h obtained with the same procedure as discussed in Section 3.2.2. The raw diagnostic metric for the complete feature set is presented in Figure 16(i) as a function of measurement number. The processed diagnostic metric and the threshold are obtained with exactly the same procedure used in Section 4.2.1. The raw diagnostic metric contains much more noise

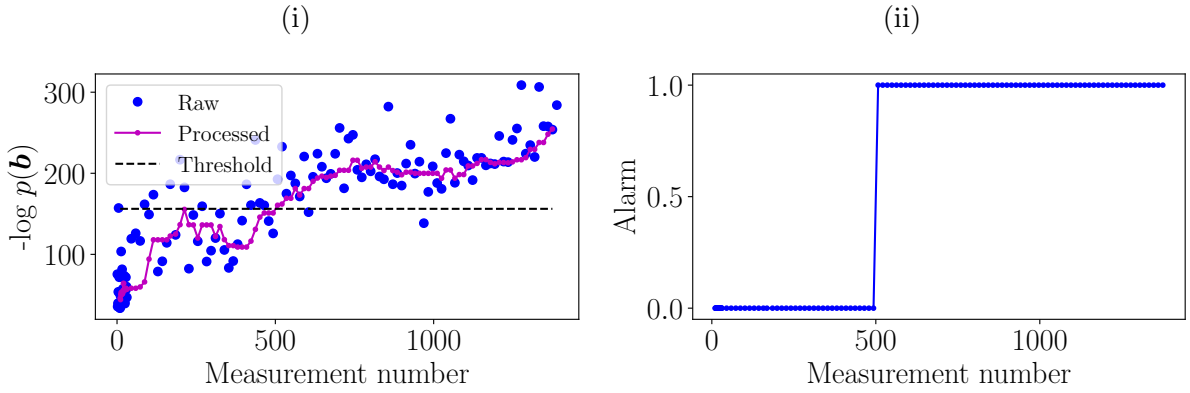

Figure 16: The raw diagnostic metric for the full dataset, $-\log p(\boldsymbol{b})$, obtained from the kernel density model applied to the experimental gearbox dataset, is shown with the processed diagnostic metric and the alarm threshold in Figure 16(i). The corresponding alarm value is presented in Figure 16(ii) as a function of measurement number.

than the raw diagnostic metric in Figure 14(i). The kernel density model is able to capture the distribution of the healthy data more accurately, which means that it is much

more sensitive to changes in the distribution of the data. The distribution of the data can for example be changed by environmental factors such as temperature variations and subtle differences in operating conditions. This means that the kernel density model is more sensitive to damage, but also inherently more sensitive to noise in the data and this is reflected in the diagnostic metric.

The alarm threshold is compared to the processed data to obtain the alarm presented in Figure 16(ii). The alarm is only triggered once at measurement 507, which is significantly later than the results in Figure 14(ii). However, this result can be misleading, because the diagnostic metric in Figure 16(i) almost results in the alarm being triggered earlier, similarly to the result of the Gaussian model in Figure 14(ii). Hence, the two models perform very similarly when comparing their diagnostic metric of the complete features.

The raw and processed marginal diagnostic metric are presented in Figure 17(i) for the gear and the pinion with the alarm threshold superimposed as well. The alarm threshold of

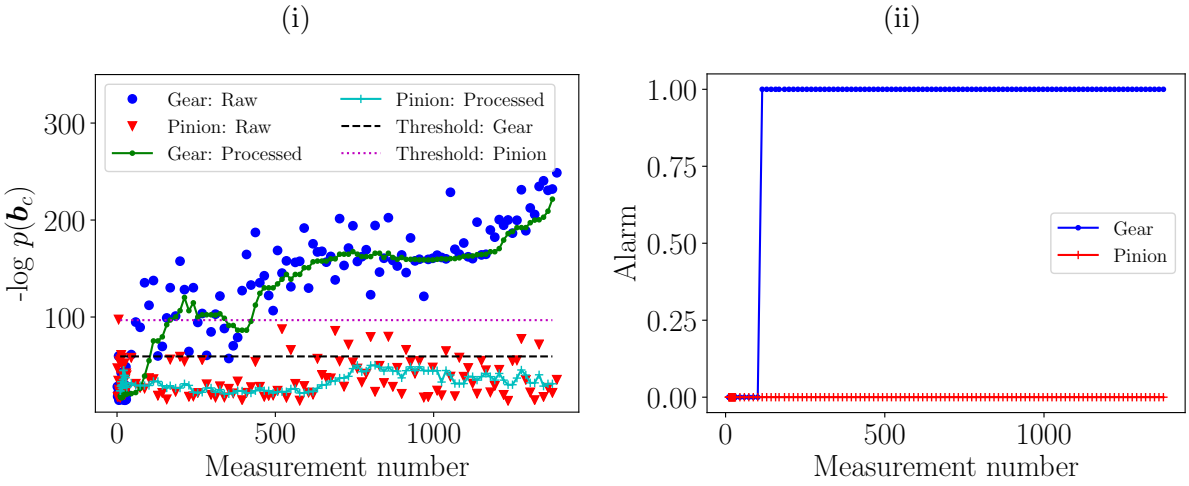

Figure 17: The raw and processed diagnostic metrics of the marginal distribution of the kernel density model and the alarm threshold is presented in Figure 17(i) for the experimental gearbox data. The alarm value, obtained by comparing the alarm threshold to the processed diagnostic metric, is presented in Figure 17(ii) for the monitored components.

the pinion is larger than the alarm threshold of the gear, because its features of the initial measurements contain more noise. It is evident from the processed diagnostic metric that significant changes are observed in the initial measurements of the gear (i.e. it has a large

gradient), which reflects the fact that the kernel density model is much more sensitive to changes in the condition, however, the larger noise levels in the gear and the pinion diagnostic metrics also indicate that the model is more sensitive to noise in the condition monitoring data as well.

The processed results in Figure 17(i) are compared to the associated alarm threshold in the same figure to obtain the alarm in Figure 17(ii). According to the alarm, the damaged gear is detected at the same measurement number as the Gaussian model, while the pinion remains healthy for the complete test. Hence, in terms of detection capabilities, the Gaussian and the kernel density models performed the same. The results in Figure 17 also emphasise the benefits of using the processed data for detection and not the raw data. The noise that is observed in the pinion in Figure 17(i) result in the threshold to be exceeded, however, the processed diagnostic metric, reduces the sensitivity of the alarms to outliers.

5. Conclusion and recommendations

In this paper, a methodology is proposed which utilises the modified improved envelope spectrum for performing feature extraction and a model of the extracted features, optimised on healthy historical data, for automatic fault detection and fault localisation under varying operating conditions. The performance of the Gaussian model and the model obtained from kernel density estimators are compared on a numerical gearbox dataset which contains bearing and gear damage, and on an experimental gearbox dataset which contains a damaged gear. In summary, the results indicate:

- The methodology is able to detect and localise the damaged component by using the negative log-likelihood of the marginal distribution of a specific characteristic as a diagnostic metric.

- The proposed features are presented and motivated based on state-of-the-art techniques in the condition monitoring field, i.e. they do not require feature engineering to be performed.

- The model obtained from a kernel density estimator is capable of estimating the probability density function of the data more accurately and this results in the

model being more sensitive to damage, but also more sensitive to spurious noise in the condition monitoring data. Its performance is also very dependent on the estimation of the hyperparameter, a task which is not easy to perform.

- The assumption made by the Gaussian model is poor for practical applications (the condition monitoring data and their features are non-Gaussian), however, it performs very well on the investigated datasets. It has the additional benefits that the model is robust to noise in the condition monitoring data and the two unknown parameters can easily be estimated with closed-form solutions.

- It is important to use the statistics of the diagnostic metrics, e.g. the median of consecutive measurements for detection, rather than the diagnostic metrics of the individual measurements.

In this paper, the monitored characteristics did not coincide, which means that it is easy to distinguish between different damaged components. However, when monitoring rotating machines with many rotating components, some overlap between the components in the order domain is inevitable. Our current suggestion is to ignore the overlapping harmonics or to assign the same label to components with the same fundamental cyclic orders. However, we do recommend that future investigations should focus on investigating the different solutions for characteristics with overlapping cyclic orders. It is also recommended that more elaborate feature modelling approaches be investigated as well. Lastly, it is also recommended that the methodology is extended for performing automatic condition recognition when historical fault data are available in addition to the healthy historical data.

Acknowledgements

S. Schmidt and P.S. Heyns gratefully acknowledge the support that was received from the Eskom Power Plant Engineering Institute (EPPEI) in the execution of the research. K.C. Gryllias gratefully acknowledges the Research Fund KU Leuven.

Appendix A. Additional information

The following procedure was used to generate each rotational speed signal $\dot{\theta}(t)$ for the numerical gearbox model:

1. Select an unscaled baseline signal, denoted $o(t)$, i.e. the range can be arbitrary. The following baseline signals are used in this paper:

$$o_1(t) = 1 - 0.1t + 0.05\sin(2.4\pi t + \phi_1^\epsilon) + 0.1\sin(1.0\pi t + \phi_2^\epsilon) \tag{A.1}$$

$$o_2(t) = \sin(0.4\pi t + \phi_3^\epsilon) + 0.2\sin(2.0\pi t + \phi_4^\epsilon) + 0.1\sin(3.0\pi t + \phi_5^\epsilon) \tag{A.2}$$

$$o_3(t) = (1 + \exp(5 - t - 10 \cdot \phi_6^\epsilon))^{-1} + 0.05\sin(2.8\pi t + \phi_7^\epsilon) \tag{A.3}$$

   where the subscript i in $o_i(t)$ denotes the operating condition profile number used in Figure 2. The phase of the rotational speed signal components, denoted ϕ_j^ϵ, is sampled from a zero mean, unit variance Gaussian distribution.

2. Linearly scale the baseline vibration signal $o_i(t)$ to have a minimum value of $\dot{\theta}_{min}$ and a maximum value of $\dot{\theta}_{max}$, where $\dot{\theta}_{min} = 2\pi\,(3 + u - 0.5)$ rad/s and $\dot{\theta}_{max} = 2\pi\,(7 + u - 0.5)$ rad/s. The variable u is a sample from a uniform distribution with domain $[0, 1]$.

This procedure was used to generate each signal shown in Figure 2, which is why the magnitude and the phase of the rotational speed signals $\dot{\theta}_i(t)$ are random.

References

[1] J. P. Salameh, S. Cauet, E. Etien, A. Sakout, L. Rambault, Gearbox condition monitoring in wind turbines: A review, Mechanical Systems and Signal Processing 111 (2018) 251–264.

[2] Y. Lin, L. Tu, H. Liu, W. Li, Fault analysis of wind turbines in China, Renewable and Sustainable Energy Reviews 55 (2016) 482–490.

[3] F. Pedro, G. Márquez, A. Mark, J. María, P. Pérez, M. Papaelias, Condition monitoring of wind turbines: Techniques and methods, Renewable Energy 46 (2012) 169–178.

[4] J. Antoni, F. Bonnardot, A. Raad, M. El Badaoui, Cyclostationary modelling of rotating machine vibration signals, Mechanical Systems and Signal Processing 18 (6) (2004) 1285–1314.

[5] J. Antoni, Cyclic spectral analysis of rolling-element bearing signals: Facts and fictions, Journal of Sound and Vibration 304 (3-5) (2007) 497–529.

[6] J. Antoni, Cyclic spectral analysis in practice, Mechanical Systems and Signal Processing 21 (2007) 597–630.

[7] C. Capdessus, M. Sidahmed, J. Lacoume, Cyclostationary processes: Application in gear faults early diagnosis, Mechanical Systems and Signal Processing 14 (3) (2000) 371–385.

[8] R. B. Randall, J. Antoni, Rolling element bearing diagnostics - A tutorial, Mechanical Systems and Signal Processing 25 (2) (2011) 485–520.

[9] D. Abboud, J. Antoni, S. Sieg-Zieba, M. Eltabach, Envelope analysis of rotating machine vibrations in variable speed conditions: A comprehensive treatment, Mechanical Systems and Signal Processing 84 (2017) 200–226.

[10] D. Abboud, S. Baudin, J. Antoni, D. Rémond, M. Eltabach, O. Sauvage, The spectral analysis of cyclo-non-stationary signals, Mechanical Systems and Signal Processing 75 (2016) 280–300.

[11] D. Abboud, J. Antoni, M. Eltabach, S. Sieg-zieba, Angle\time cyclostationarity for the analysis of rolling element bearing vibrations, Measurement 75 (2015) 29–39.

[12] S. Schmidt, P. S. Heyns, K. C. Gryllias, A probabilistic novelty detection methodology based on the order-frequency spectral coherence, International conference on Condition Monitoring of Machinery in Non-stationary Operations (CMMNO2018).

[13] C. M. Bishop, Pattern recognition and machine learning, Springer, 2006.

[14] D. Abboud, J. Antoni, S. Sieg-Zieba, M. Eltabach, Deterministic-random separation in nonstationary regime, Journal of Sound and Vibration 362 (2016) 305–326.

[15] J. Antoni, G. Xin, N. Hamzaoui, Fast computation of the spectral correlation, Mechanical Systems and Signal Processing 92 (2017) 248–277.

[16] P. Borghesani, J. Antoni, A faster algorithm for the calculation of the fast spectral correlation, Mechanical Systems and Signal Processing 111 (2018) 113–118.

[17] S. Khan, T. Yairi, A review on the application of deep learning in system health management, Mechanical Systems and Signal Processing 107 (2018) 241–265.

[18] D. Abboud, J. Antoni, Order-frequency analysis of machine signals, Mechanical Systems and Signal Processing 87 (October 2016) (2017) 229–258.

[19] R. B. Randall, J. Antoni, S. Chobsaard, The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals, Mechanical Systems and Signal Processing 15 (5) (2001) 945–962.

[20] S. Schmidt, P. S. Heyns, K. C. Gryllias, A discrepancy analysis methodology for rolling element bearing diagnostics under variable speed conditions, Mechanical Systems and Signal Processing 116 (2019) 40–61.

[21] C. M. Bishop, Novelty detection and neural network validation, IEE Proceedings - Vision, Image, and Signal Processing 141 (4) (1994) 217.

[22] R. Zimroz, W. Bartelmus, T. Barszcz, J. Urbanek, Diagnostics of bearings in presence of strong operating conditions non-stationarity - A procedure of load-dependent features processing with application to wind turbine bearings, Mechanical Systems and Signal Processing 46 (1) (2014) 16–27.

[23] W. Bartelmus, R. Zimroz, A new feature for monitoring the condition of gearboxes in non-stationary operating conditions, Mechanical Systems and Signal Processing 23 (5) (2009) 1528–1534.

[24] J. Urbanek, T. Barszcz, R. Zimroz, J. Antoni, Application of averaged instantaneous power spectrum for diagnostics of machinery operating under non-stationary operational conditions, Measurement 45 (7) (2012) 1782–1791.