Mechanical Systems and Signal Processing 205 (2023) 110833

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

Data-driven segmentation of long term condition monitoring data in the presence of heavy-tailed distributed noise with finite-variance

Hamid Shiri [a,*], Pawel Zimroz [a], Jacek Wodecki [a], Agnieszka Wyłomańska [b], Radoslaw Zimroz [a]

[a] Faculty of Geoengineering, Mining and Geology, Wroclaw University of Science and Technology, Na Grobli 15, 50-421 Wroclaw, Poland
[b] Faculty of Pure and Applied Mathematics, Hugo Steinhaus Center, Wroclaw University of Science and Technology, Wyspianskiego 27, 50-370 Wroclaw, Poland

ARTICLE INFO

Communicated by P. Borghesani

Keywords:
Long-term data
Segmentation
Robust methods
Non-gaussian noise
Changing points

ABSTRACT

Machinery condition prognosis systems use long-term historical data to predict the remaining useful life (RUL). One of the critical steps to reach this purpose is to segment long-term data into two or several degradation stages (healthy, unhealthy, critical stage). Finding a changing point between stages may be a crucial preliminary task for further prediction of degradation process. However, finding the accurate partition into two or more stages is a challenging task in actual application when noise inherent in the observed process exhibits non-Gaussian characteristics. In this paper, a framework for data-driven segmentation is presented for prognosis of machinery long-term data in presence of heavy-tailed distributed noise with finite variance. It is assumed that three different stages are inherent in degradation process and each segment of data follows a specific trend (constant, linear, exponential or polynomial). At first, data is divided into three parts. Trend functions are fitted to the data by using robust regression method, and cumulative error is calculated. This process is done iteratively for all possible partitions into three intervals to find the segmentation which minimizes the error. The framework has been tested via empirical analysis of estimators of the changing points obtained in Monte Carlo simulations. Also, discussed approaches are applied to the real data. In such measurement, data that are commonly available (in condition monitoring systems) is aggregated from the raw signal and sampled at long intervals. Finally, effectiveness of the segmentation results is assessed by comparing them with envelope frequency analysis of raw signal to confirm the fact that detected changing points coincide with start time of the fault in the machine or not.

1. Introduction

Prognostics and Health Management (PHM) is one of the significant tasks in condition-based maintenance (CBM). The CBM systems are developed to assess the machine's condition by collecting massive amounts of data during the operation of machines. PHM is usually composed of four different processes: 1. Data acquisition is defined as the process of recording and saving different types of monitoring data from various sensors mounted on the monitored machine. 2. Construction of a health index (HI): Generally, a health index is defined as the set of statistical and frequency domain-based features extracted from raw signals to describe health condition of the machinery. It should be noted that selecting an appropriate health index that accurately describes degradation is crucial for subsequent procedures. Extensive literature exists on this topic, with numerous papers specifically addressing this area

* Corresponding author.
E-mail address: hamid.shiri@pwr.edu.pl (H. Shiri).

https://doi.org/10.1016/j.ymssp.2023.110833
Received 11 April 2023; Received in revised form 27 August 2023; Accepted 28 September 2023
Available online 11 October 2023

of research. However, due to the scope of this paper, we will refrain from providing an in-depth description of this aspect. 3. Health stage (HS) evaluation and 4. Predicting RUL: HS evaluation is taken as an expression of fault detection in the PHM community [1]. Nevertheless, their assignments are different from each other. Fault diagnosis is determined by the pattern and intensity of the fault of the machine. However, the primary purpose of HS evaluation is to divide continuous degradation process into several health stages due to varying trends of HI. Also, it is hard and dubious to predict RUL during the healthy stage. RUL estimation should start from the unhealthy stage's start time, known as the first predicting time (FPT), due to the fact that in the healthy stage observations, there is no valuable information about the unhealthy stage and its degradation trend. Therefore, instead of trying too hard to employ complex model, such as the model with switching regimes or non-linear trends to predict RUL, breaking HI into different HS and trying to predict RUL from the last stage may be more accessible and evaluative. Commonly, for HS evaluation, the health index is compared to a limit value that usually is provided by the manufacturer (threshold corresponding to change from "Good Condition" (healthy stage) to "Warning" (degradation stage) and "Warning" to "Alarm" stage (critical stage). Unfortunately, we do not know the limit values or desired useful life in many cases, which is often present in raw materials industry [2], especially when the machine is unique. In addition, note that most of these thresholds are introduced by manufacturing industries and are defined based on particular working and environmental conditions standards that may not be valid when conditions change.

In recent years, several papers have been published in the literature to detect changing points in time series in different areas such as financial, medical, and meteorology [3–8]. Also, it should be noted that this process is known as regime-switching point detection or signal segmentation in other communities. Signal segmentation is repeatedly employed in signal processing applications to separate original data into homogeneous segments or extract the pattern. In related studies, qualitative trend analysis (QTA) has been employed as a method for achieving the same objective. QTA involves analyzing and interpreting qualitative patterns and trends in the data to identify stage borders [4,5,9]. However, QTA as an approach which takes into account trend only. In our case, the considered problem is much more general. Gasior et al. [10] used segmentation for shock extraction in sieving screen vibrations. Kucharczyk et al. [11] used stochastic modeling for seismic signal segmentation. Grzesiek et al. [12] proposed a methodology to detect regime-changing when regime A smoothly transforms into regime B. Also, a few papers have been published in the PHM community based on dividing HI into different HS. Some of these articles divide HI into two different stages. Alkan et al. [13] presented a methodology for diagnosis of electromechanical systems faults based on the variance-sensitive adaptive alarm threshold and principal component analysis (PCA). Fink et al. [14] explained the prediction of RUL as a two-stage classification to detect the machine's condition after the defined time interval. Schlechtingen et al. [15] compared three different model-based approaches for two-stage division of wind turbines, i.e., a regression-based model and two artificial neural networks (ANN) based models. Hu et al. [16] used a one-class support vector machine and a Gaussian threshold model for condition monitoring of turbo-pumps. The two-stage division is only valuable for cases where the degradation trends of machinery in the unhealthy stage are consistent and can be represented utilizing a single degradation model. Nonetheless, due to variations of fault patterns or operational conditions, the degradation trends of machinery may change. According to this circumstance, the unhealthy stage should be split into various stages based on different degradation trends. A few of these research divided HI and spectra into several stages, with changing points detection approaches. For instance, Kimotho et al. [17] segmented degradation trend into five stages due to variation of frequency amplitude in power spectra density. Sutrisno et al. [18] split the bearing degradation process into different stages by employing anomaly detection of frequency spectra. Hu et al. [19], separated HI into four stages using changing points of confidence levels. Methods based on machine learning techniques are frequently used for long-term data analysis for classification, fault detection, prognosis, etc. [2,20–23]. In this case, unsupervised classification, i.e. clustering, is required because every machine works under different conditions and labeling data is not an easy task. Hence, the clustering approach has a huge potential to divide long-term data into several regimes. The authors of [21] proposed the Long Short Term Memory (LSTM) network with a clustering method for multi-stage predicting RUL. Jaskaran Singh et al. [22] introduced an adaptive data-driven model-based approach to detect regime-changing points using K-means clustering. Also, discrete state transition models such as HMMs [24–26] and dynamic state-space models [27–29] are frequently employed to segment degradation processes into multiple stages. However, most of the mentioned research was performed under the assumption that observation noise has Gaussian distribution, while in most actual applications, especially in PHM area it is not a proper assumption. In many cases, impulsive noise can be observed, which is properly described with a heavy-tailed distribution, meaning that marginal events are likely to occur. This may happen for several reasons. Wind inflows are one of possible sources of impulsive noise in wind turbine [30–32], ore falling on devices in the mining environment is another [33,34].

Due to specific random character of degradation process observed in HI data, classical algorithms known from the literature may fail to properly perform the segmentation procedure [35–37]. To solve this problem, in the article, we propose a novel stochastic data-driven approach that leverages whole historical data to robustly identify stage borders, specifically the transitions between the healthy, degradation, and critical stages. Our method is designed to handle the presence of non-Gaussian noise with finite variance, which commonly occurs when the machine undergoes condition changes. It is considered that degradation process is composed of three stages and each stage of the data follows a specific trend (constant, linear and exponential or polynomial). First, the long-term historical data is divided into three parts. Trend functions are fitted to the data using the regression method and cumulative error is calculated. This process is done iteratively for all possible partitions into three intervals to find the changing points, which minimizes the error. This procedure is repeated using different regression methods including robust methods such as the least absolute error and fitting the Student's t distribution. It should be noted that the Student's t distribution was chosen because it is both heavy-tailed (unlike Gaussian distribution) and at the same time has finite variance (for considered range of degree of freedom) making it easier to fit than some other theoretically well-established distributions. Also, in our previous paper, we showed that this distribution can describe a random part as proper [38]. Furthermore, for better comparison and to check if non-Gaussian noise assumption is

necessary, along with the mentioned approach, we have included two hidden Markov model-based methods having Gaussian noise assumption (these two methods use a relatively different structure for segmentation). We also presented a waterfall plot of the short-time envelope spectrum calculated for available raw vibration data for the real data sets. Such a 3D plot demonstrates how envelope spectrum changes and allow us to visually validate whether the detected changing points are coinciding with the time the fault in the machine occurs or not.

The main contribution of this paper is summarized as follows:

- We illustrated the possibility of segmenting long-term condition monitoring data into three segments based on global trends (deterministic parts) with constant (for healthy stage), linear (for degradation stage), and exponential or polynomial trend (critical stage).
- We demonstrated and highlighted the effect of noise (noise with Gaussian distribution as a standard assumption in literature and noise with non-Gaussian characteristics) to detect mentioned stages. We have highlighted importance of this issue and we state there is lack of sources that analyze the effect of noise to detect any of the stages in the literature for machinery prognosis.
- We used several conventional statistical approaches (robust and non-robust) to fit a proposed model with three stages in the presence of different noise levels with time-varying scale of noise to show the performance of it to do this task and make comparisons between them.
- Additionally, to validate the results in a real case, we applied all the methods to three benchmark data sets in the prognosis area, and we evaluated the results with condition monitoring techniques to see if the model with three stages has any correlation with the nature of growing faults in the bearings and which methods perform better in actual area.

Finally, it should be highlighted, that segmentation is a crucial step in prognosis. Many authors assume they know this is second or third regime and then they do prognosis. But if we do not know what regime we have at the moment, situation is extremely difficult: if we are in regime 1 - there is no sense to do prognosis. (No one wants to forecast constant value). The difference between regimes 2 and 3 is significant, there are different models of HI behavior. In the paper we showed that to achieve information about current regime is not trivial, even if one has all historical data. Additional trouble is related to the presence of non-Gaussian noise. This is also not clear for many engineers and scientists.

The paper is structured as follows: after the introduction, in Section 2, we reviewed the most common long-term data model and described the model used in this paper. In Section 3, we have reviewed theoretical and mathematical aspects of the methods used in this paper. In Section 4, we generated data based on Section 2 and employed the proposed structure to segment long-term data in the presence of Gaussian and non-Gaussian noise. Finally, in Section 5 the results of applying the proposed approach to three benchmark data sets are presented with an indication of all intermediate steps, and in Section 6 the conclusions are formed.

2. Long term data model

Model-based degradation process analysis is a common approach in the PHM community. These models can be described by various deterministic trends (linear, polynomial, exponential, and constant); see Fig. 1, various noise distributions, and structures of dependence for random components; see Figs. 2, 3.

According to pre-knowledge of degradation process, the model can be selected; for example, linear model is frequently used to describe the drilling degradation process [39,40], see panel (b) Fig. 1. Meanwhile, exponential degradation trend is usually used for expressing the battery degradation process [41,42] (but with a negative slope), see panel (c) in Fig. 1. To describe degradation process of bearings, gears or shafts, different kinds of model composed from various deterministic trends may be used [43,44] (see Fig. 1, panel (d), (e), (f)).

In this paper, we assume that degradation process has complex trends, such as those presented in Fig. 3. This model consists of three stages. Firstly, degradation process is centered around some constant level that corresponds to healthy stage. The second stage is made up with a linear function corresponding to the degradation stage. The last part has either an exponential or a polynomial trend that is referred to as the critical stage. The two moments of transition from one stage to another we will denote as changing points: CP1 (is a border between healthy stage with constant trend and degradation stage with linear trend) and CP2 (is a border between degradation stage with linear trend and critical stage with exponential or polynomial trend).

Also, this model can include different kinds of noise distribution covering degradation process. In most of the research in this area, it is assumed the distribution of noise is Gaussian, see Fig. 2 [45], while this assumption is not proper when the machine works in harsh environments [38]. To achieve a model that better reflects real condition, in our previous work [38] we developed a simulation framework that produces degradation process observations; an exemplary trajectory is shown in Fig. 3. In this framework, along with the trend of degradation process, the distribution of noise and its scale (square root of variance in the Gaussian case) are also changing in each stage. Main characteristics of our framework are summarized in Table 1. In this paper, we assume finite variance for non-Gaussian distribution.

3. Methodology

We assume that the long-term data of health index (HI) is composed of three distinct stages exposing significantly different behavior. Based on this assumption, our aim is to divide the time series into these three sections using robust mathematical regression methods, achieving the best fit to one of the specified models. All of the methods that we discuss, assume that the health index is a

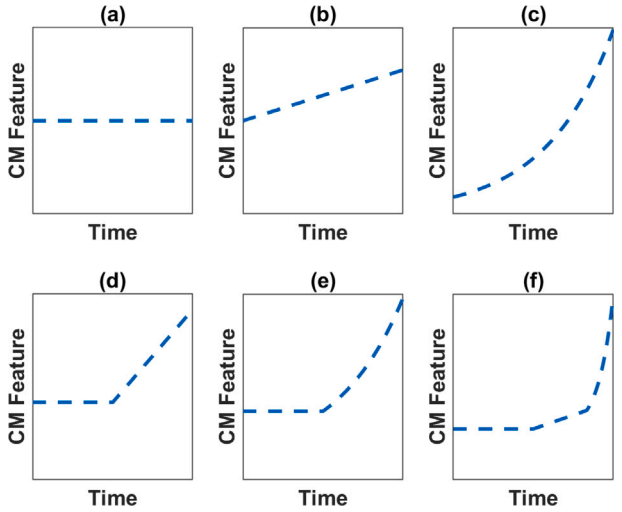


Fig. 1. Different types of HI variations and degradation models: (a) good condition (constant trend), (b) good to gradual wear (linear trend), (c) exponential trend, (d) good to accelerated wear (linear trend), (e) good to accelerated wear (exponential trend), (f) three stages model (good, linear progress and exponential (polynomial) progress of degradation) [38].

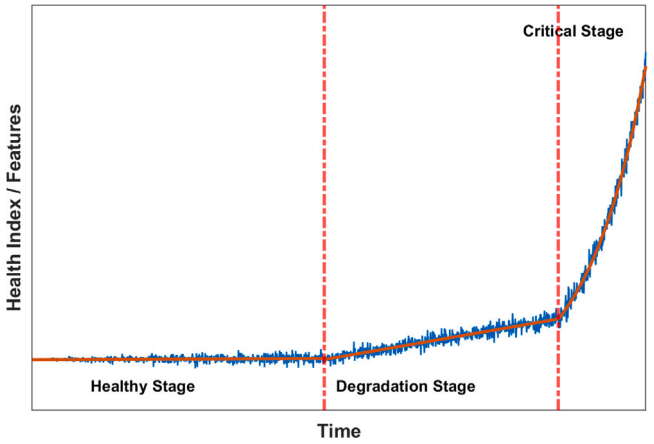


Fig. 2. Three stage degradation model [46].

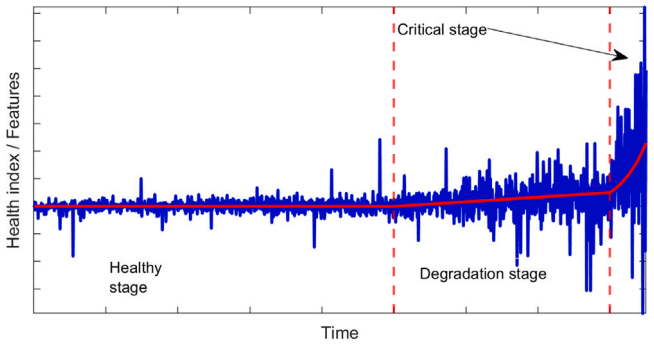


Fig. 3. The proposed long-term degradation model with 3 stages [38].

stochastic process (its observations we denote by X_t where t is the discrete-time index) and that its trend is a function of time and a set of parameters. Thus, the HI observation process consists of time-changing trend (expected value of X_t, which we assume to be

Table 1
Main characteristics of data for the three stages indicated in Fig. 3 [38].

	Stage 1	Stage 2	Stage 3
Trend	Constant	Linear	Exponential or polynomial
Scale	Nearly constant	Linearly growing	Lin. or exp. growing
Random component	Gaussian/non-Gaussian	Gaussian/non-Gaussian	Gaussian/non-Gaussian

finite) and residual r_t, hence we have the following.

$$X_t(\Theta) = \widehat{X}_t(\Theta) + r_t(\Theta), \qquad \widehat{X}_t(\Theta) = E(X_t(\Theta)). \tag{1}$$

By Θ in this section, we denote the complete set of parameters to be fitted to the data. It may include the changing points (see previous section) and the parameters of a probability distribution if a method employs any. We now present the methods in two groups.

3.1. Piece-wise regression models

This class of methods allows for division of the data into three segments, each having a different form of the model. We employ three functional components: constant, linear, and exponential as below:

$$\widehat{X}_t(\Theta) = \begin{cases} f_1(t,\theta_1) = \theta_1, & 0 < t \le \tau_1, \\ f_2(t,\theta_2,\theta_3) = \theta_2 t + \theta_3, & \tau_1 < t \le \tau_2, \\ f_3(t,\theta_4,\theta_5,\theta_6) = \theta_4 \exp(\theta_5 t) + \theta_6, & \tau_2 < t \le N, \end{cases} \tag{2}$$

where the variables τ_1 and τ_2 stand for CP1 and CP2, the number N is signal length, and $\theta_1, \dots, \theta_6$ are constants (which need to be fitted) parameters used to model deterministic parts of HI.

3.1.1. Ordinary least squares (OLS)

The most basic approach is based on minimizing sum of squares of residuals of the model. The procedure designed to find optimal set of parameters iterates in two loops through possible divisions into three windows (which allows us to find optimal τ_1 and τ_2).

$$\widehat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \sum_{t=1}^{\tau_1 - 1} (X_t - f_1(t,\theta_1))^2 + \sum_{t=\tau_1}^{\tau_2 - 1} (X_t - f_2(t,\theta_2,\theta_3))^2 + \sum_{t=\tau_2}^{N} (X_t - f_3(t,\theta_4,\theta_5,\theta_6))^2. \tag{3}$$

In the case of observations with Gaussian noise, given each division, optimal parameters can be equivalently estimated using Maximum Likelihood Estimation (MLE) with Gaussian distributed residuals.

3.1.2. Dynamic programming segmentation

In this method, the optimal segmentation of the time series of observations $X_{t=1,\dots,N}$ is done by fitting model parameters in terms of Gaussian distribution MLE by using optimized dynamic programming. The computation cost of this model is high based on using dynamic programming. For more details on this method, please see following Refs. [47,48].

3.1.3. Iteratively reweighted least squares (IRLS)

The method of iteratively reweighted least squares (IRLS) is a modification of the OLS method where the weights are assigned to error terms according to changing variance which we assume to be finite (for each time t the weight $w_t = 1/\sigma_t$ is reciprocal of the corresponding standard deviation σ_t of the error term of the observation). The weights are modified in an iterative way; thus, in the jth step, the vector of weights $w_t^{(j)}, t = 1, \dots, N$ is applied, resulting in updating of the piecewise regression parameters:

$$\widehat{\Theta}^{(j)} = \underset{\Theta}{\operatorname{argmin}} \sum_{t=1}^{\tau_1 - 1} w_t^{(j)} (X_t - f_1(t,\theta_1^{(j-1)}))^2 + \sum_{t=\tau_1}^{\tau_2 - 1} w_t^{(j)} (X_t - f_2(t,\theta_2^{(j-1)},\theta_3^{(j-1)}))^2 + \sum_{t=\tau_2}^{N} w_t^{(j)} (X_t - f_3(t,\theta_4^{(j-1)},\theta_5^{(j-1)},\theta_6^{(j-1)}))^2, \tag{4}$$

where $\widehat{\Theta}^{(j)}$ is the current values set of Θ in the jth iteration of IRLS algorithm. Among many different iterative methods of finding proper weights, we used a method that applies to model's residuals the Tukey biweight function $\phi(x) = x(1-x^2)^2$ (as defined in [49]) implemented in Matlab robustfit() function. Weights are updated using the formula [50]:

$$w_t^{(j)} = \frac{\phi(\varepsilon_t^{(j)})}{\varepsilon_t^{(j)}}, \tag{5}$$

where $\varepsilon_t^{(j)}$ is residual of the model divided by the current estimate of standard deviation term, that is:

$$\varepsilon_t^{(j)} = \frac{r_t}{\sigma_t^{(j)}}. \tag{6}$$

3.1.4. Least absolute error (LAE)

The sum of absolute deviations is used as a cost function instead of a sum of squares. It is a widely known alternative to OLS method, making adjustment of the regression parameters more robust, even in the case of infinite variance [51]. The optimization procedure that includes LAE cost function with changing points is defined as follows:

$$\widehat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \sum_{t=1}^{\tau_1-1} |X_t - f_1(t,\theta_1)| + \sum_{t=\tau_1}^{\tau_2-1} |X_t - f_2(t,\theta_2,\theta_3)| + \sum_{t=\tau_2}^{N} |X_t - f_3(t,\theta_4,\theta_5,\theta_6)|. \tag{7}$$

The same as in OLS method we iterate through all possible divisions into three stages and in that way we find τ_1 and τ_2. This method is beneficial when the errors have rather a heavy-tailed than Gaussian distribution. Namely, given each division into three windows, it can be equivalently stated as applying MLE estimation with Laplace distributed residuals.

3.1.5. Student's t distribution estimation (ST)

In this method, residuals of the model in each degradation stage are assumed to have a scaled Student's t distribution, that is:

$$r_t = \begin{cases} \sigma_1 u_t, & 0 < t \le \tau_1, \\ \sigma_2 u_t, & \tau_1 < t \le \tau_2, \\ \sigma_3 u_t, & \tau_2 < t \le N, \end{cases} \tag{8}$$

where $\{u_t\}_{t=1,\dots,N}$ is a series of independent random variables from Student's t distribution with ν_i degrees of freedom for the i -th segment ($i = 1, 2, 3$). This enables heavy tails in the error distribution and makes fitting more robust allowing proper parametrization of non-Gaussian observation data. Estimation procedure is based on MLE technique applied to each possible division into three stages. The optimal set of parameters is found to fulfill the following:

$$\widehat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \sum_{t=1}^{\tau_1-1} \log\Big(p_{\nu_1}\Big(\frac{X_t - f_1(t,\theta_1)}{\sigma_1}\Big)\Big) + \sum_{t=\tau_1}^{\tau_2-1} \log\Big(p_{\nu_2}\Big(\frac{X_t - f_2(t,\theta_2,\theta_3)}{\sigma_2}\Big)\Big) + \sum_{t=\tau_2}^{N} \log\Big(p_{\nu_3}\Big(\frac{X_t - f_3(t,\theta_4,\theta_5,\theta_6)}{\sigma_3}\Big)\Big), \tag{9}$$

where $p_\nu(\cdot)$ is the probability density function (PDF) of Student's t distribution given by the formula [52]:

$$p_\nu(y) = \frac{(1 + \frac{y^2}{\nu})^{-\frac{\nu+1}{2}}}{B(\frac{\nu}{2}, \frac{1}{2})\sqrt{\nu}}, \qquad y \in \mathbb{R}, \tag{10}$$

where $B(\cdot,\cdot)$ is the beta function. The parameter ν is called number of degrees of freedom. We restrict parameter ν to be greater than 2, thus variance of the underlying random variable is finite [52].

Additionally, it is worth noting that optimal values for parameters of the methods described in this subsection (namely OLS, IRLS, LAE and ST) are determined based on unconstrained multivariate optimization using derivative-free approach (we use fminsearch function in Matlab). Also, we should again emphasize the fact that in the proposed approach described with Eq. (1), the three stages are only assumed to exist but the values of CPx are not fixed at the initialization of the parameters fitting procedure. Due to Eqs. (3) to (10), all the parameters of the model can be fitted, including also τ_1 and τ_2.

3.2. Hidden Markov models and expectation–maximization algorithm

For a shorter notation, in this subsection, we will denote the process of observations as $X = \{X_t\}_{t=1,\dots,N}$. Hidden Markov models are the models in which to describe the evolution of observations process X, an additional unobserved process (hidden states) $z = \{z_t\}_{t=1,\dots,N}$ is introduced so that the whole system has Markov property. As hidden states are unknown, the full likelihood of a system $P(X, z|\Theta)$ cannot be calculated directly to obtain the parameters with MLE method. Standard iterative method that can be used instead is called Expectation–Maximization (EM) algorithm. First, some initial values of $\Theta^{(0)}$ are set. Next, EM procedure is based on repeated steps of calculating expectation of the likelihood given the distribution of z conditioned with X and $\Theta = \Theta^{(u)}$ and reassigning Θ to new values that maximize conditional expectation. The uth step can be described with the updating formula:

$$\Theta^{(u+1)} = \underset{\Theta}{\operatorname{argmax}}\, E_{z|X,\Theta=\Theta^{(u)}}\Big(P(X,z)|\Theta\Big), \tag{11}$$

where the letter E stands for conditional expectation. We consider two different models from this class. However, they share the same equation for evolution of X given z, which is a polynomial (of degree p) regression:

$$X_t = \sum_{i=0}^{p} \beta_{i,z_k} t^i + \sigma_{z_t} \varepsilon_t, \tag{12}$$

where $t = 1, \dots, N$, the constants β_{0,z_k} to β_{p,z_k} are polynomial coefficients in the state (degradation stage) z_k, and ε_t is Gaussian white noise. The set of $p+1$ polynomial coefficients depends on the current stage z_t. The hidden states can be interpreted as degradation stages indexed from 1 to K in the proper order from healthy state to full degradation (thus K denotes number of distinct degradation stages). In this paper, according to the framework described in Section 2, we put K = 3.

In the HMM approach, CP1 and CP2 are not included in Θ. The result of fitting HMM model to the data contains probabilities for each t of the current stage z_t being equal to $1, \dots, K$. The state with highest probability indicates in which degradation state the observed machine is at time t, and thus we obtain the division of degradation process into three stages.

3.2.1. Hidden Markov model regression (HMMR)

This approach employs a mixture of polynomial regressions handled by hidden states of a discrete-time Markov chain. Changing of the stages is therefore fully described by the initial distribution and the one-step transition matrix. Significant restrictions are therefore put to the transition probabilities to better reflect nature of the process, thus for all $l = 1, \dots, K$ we have:

$$P\Big(z_{t+1} = k | z_t = l\Big) = 0, \quad \text{for } k \notin \{l, l+1\} \tag{13}$$

and generally nonzero otherwise. The parameters are estimated through Baum–Welch algorithm as described in [47].

3.2.2. Hidden logistic process (HLP)

In this method, the process of unobserved switching stages affecting the set of polynomial regression coefficients is modeled with a multinomial logistic regression model. Distribution of the current stage z_t in this model is assumed to be as follows:

$$\pi_{t,k}(\beta) = P(z_t = k|\beta) = \frac{exp(\sum_{i=0}^{p} \beta_{ki} t^i)}{\sum_{l=1}^{K} exp(\sum_{i=0}^{p} \beta_{li} t^i)}, \qquad \text{for } k = 1, \dots, K, \tag{14}$$

where $\beta = [\beta_{ki}]_{k=1,\dots,K,\, i=0,\dots,p}$ is matrix containing $p+1$ regression parameters for all K degradation stages. As z is not observed, total log-likelihood of the model must include Gaussian distribution of X_t given z_t. In terms of probability density functions, we have (see [47]):

$$p(X|\Theta) = \sum_{t=1}^{N} \log \sum_{k=1}^{K} \pi_{t,k}(\beta) p_{\Theta}(X_t | z_t = k), \tag{15}$$

where the Gaussian PDF denoted with $p_{\Theta}(X_t|z_t = k)$ comes from Eq. (12). The resulting model is a special case of time-heterogeneous Gaussian mixture model. A dedicated EM for this approach is described in [47]. As is shown there, in a single iteration of EM the expectation term which is a function of Θ (compare Eq. (11)) can be separated into two terms, one dependent on β and the other dependent on σ and β. Thus, fitting of these two sets of parameters can be done independently. Matrix β is estimated utilizing a multi-class iterative reweighted least squares (IRLS) approach.

4. Analysis of simulated data

Based on the assumption of Section 2, the model for long-term data is used to generate HI observations. After the mentioned methodologies are employed to segment the data into three stages in the presence of Gaussian noise and non-Gaussian noise. Additionally, to show performance of the methodology, results were compared together.

4.1. Model description

Considering the model discussed in Section 2, we propose following model to generate $HI(t)$ that will be used in the simulation study:

$$HI(t) = R(t) + D(t), \tag{16}$$

where $R(t)$ and $D(t)$ are, respectively, random and deterministic components. Both these parts consist of three stages, denoted as stage 1, stage 2 and stage 3, which are related to three considered underlying stages (healthy stage/degradation stage/critical stage) and determine behavior of the process with regard to both trend and noise's scale (variance in Gaussian distribution). Let us assume that we have a sample signal $HI(1), \dots, HI(N)$. The changing point between stages 1 and 2 is denoted by τ_1, and the changing point between stages 2 and 3 is τ_2, where $1 < \tau_1 < \tau_2 < N$. In other words, we can divide the signal segment by segment, so that the sequence $HI(1), \dots, HI(\tau_1)$ corresponds to stage 1, the sequence $HI(\tau_1 + 1), \dots, HI(\tau_2)$ corresponds to stage 2 and the sequence $HI(\tau_2 + 1), \dots, HI(N)$ corresponds to stage 3.

In the simulation study presented in the next section, we assume two distributions for the random component, namely Gaussian and Student's t. The aim of using both Gaussian and Student's t (being an example of a heavy-tailed distribution) in the simulation is to provide a comparison between them for condition monitoring application, as non-Gaussian behavior is widely present in real scenarios. The random component corresponding to $R(t)$ is constructed in the following way:

$$R(t) = SC(t)\widetilde{R}(t), \tag{17}$$

where function $SC(t)$ represents the time-changing scale and $\widetilde{R}(t)$ is a series of independent identically distributed random variables. For discussion and interpretation of the components of model (17) we refer the Reader to our previous article [38]. For Gaussian distribution, for each t we put $\widetilde{R}(t) \sim \mathcal{N}(0,1)$ and for Student's t distribution case $\widetilde{R}(t) \sim t(\nu)$. For simplicity, we assume that the distribution in each stage is the same, however, as it was mentioned in Section 2, in practice, it may be different for different stages.

As it was mentioned, its behavior is different for each stage, namely, we assume that in stage 1 the scale grows linearly, from σ_1 to σ_2 (where both values are relatively close to each other), then in stage 2 it also increases linearly, from σ_2 to σ_3, and finally, in stage 3 it grows exponentially, from σ_3 to σ_4. The function $SC(t)$ is defined as follows:

$$SC(t) = \begin{cases} a_1 t + b_1 & 0 < t \le \tau_1, \\ a_2 t + b_2 & \tau_1 < t \le \tau_2, \\ a_3 \exp(b_3 t) & \tau_2 < t \le N, \end{cases} \tag{18}$$

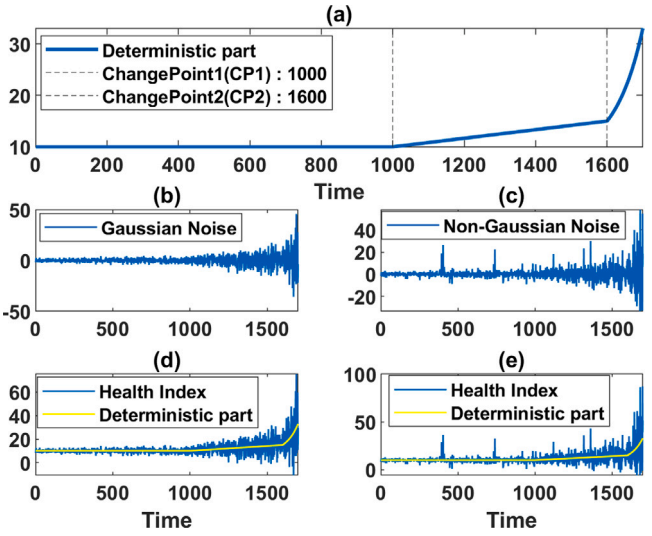


Fig. 4. Simulation model: (a) deterministic part with marked changing points, (b) Gaussian noise, (c) non-Gaussian noise with $\nu = 3$, (d) health index in presence of Gaussian noise, (e) health index in presence of non-Gaussian noise.

where constants a_1, b_1, a_2, b_2, a_3, b_3 are derived in such a way that $SC(1) = \sigma_1$, $SC(\tau_1) = \sigma_2$, $SC(\tau_2) = \sigma_3$ and $SC(N) = \sigma_4$ [38].

Behavior of the deterministic component $D(t)$ in Eq. (16) has different nature for different stages. Let us recall that in stage 1 it is at fixed constant level, denoted here as c_1. Then, in stages 2 and 3 we consider linear and exponential functions, respectively, with the same growth parameters as for the corresponding stages of scale function $SC(t)$. Moreover, we assume the function $D(t)$ has no discontinuities in stage changing points τ_1 and τ_2. Under these assumptions, the deterministic term of the signal has the form

$$D(t) = \begin{cases} c_1 & 0 < t \le \tau_1, \\ a_2 t + c_2 & \tau_1 < t \le \tau_2, \\ a_3 \exp(b_3 t) + c_3 & \tau_2 < t \le N, \end{cases} \tag{19}$$

where c_2 and c_3 are derived in such a way that $D(t)$ is a continuous function. In panel (a) in Fig. 4 we present deterministic component $D(t)$ and scale function $SC(t)$ for the following values of parameters: $\tau_1 = 1000$, $\tau_2 = 1600$, $N = 1700$, $\sigma_1 = 1$, $\sigma_2 = 2$, $\sigma_3 = 7$, $\sigma_4 = 25$ and $c_1 = 10$. Here, we assumed four additional values of ν- parameter, namely $\nu \in \{2.1, 3, 5, 10\}$. As can be seen, for the Gaussian distributed signal we do not observe outlying observations, see panel (b) in Fig. 4, while for the non-Gaussian heavy-tailed case, large impulses occur in the data, see panel (c) in Fig. 4. The smaller the ν, the higher impulses may occur in the signal.

It should be noted that the model used for the simulation is different from that used in the segmentation methods. A model used for piecewise regression methods (see Eq. (2)) is a simplified version of the simulated model presented in this section neglecting the scaling factor $SC(t)$, i.e. changing it to a constant. The reason for such is to investigate how complex behavior may be modeled with a simplified approach.

4.2. Analysis for Gaussian noise

In this subsection, the proposed methodology is applied to data generated by the proposed model assuming that noise term has Gaussian distribution $\widetilde{R}(t) \sim \mathcal{N}(0, 1)$. Segmentation results for simulated data are presented together for all of the selected methods. Due to the simulation procedure, at Time = 1000 healthy stage is going to degradation stage (CP1 which is the meaning of τ_1), and at Time = 1600 is CP2 (which is the meaning of τ_2).

In Fig. 5 we can observe an exemplary simulated trajectory along with estimated changing points CP1 and CP2 resulting from each of the methods. As we can see, in this case, where all stages include Gaussian noise, the selected methods could detect CP1 with good precision. CP1 is detected by most of the methods such as ST, LAE, and IRLS precisely, while rest of the methods also discovered this point somewhere around Time = 1000 and Time = 1088, which is acceptable except HLP method that discovered this point at Time = 882 happens earlier than expected time. According to this fact, transition between the two last stages is significant; CP2 was found with most methods except IRLS, OLS, and HLP, which could not detect this point as an appropriate point.

Estimation procedure was repeated for 100 simulations from the same model. In Fig. 6 the estimation results are visualized using box plots. Horizontal red lines are at the level of the mean estimation results and the level of the gray dashed line is the true changing point value. The vertical length of the blue boxes corresponds to IQR (an inter-quartile range which is the range between Q1 and Q3). In addition, outliers can be observed. Estimation of CP1 and CP2, according to HMMR, HLP and DPS, has low variance (it can be inferred from small value of IQR), however, in the case of CP1 the result is biased to the right from the true value of

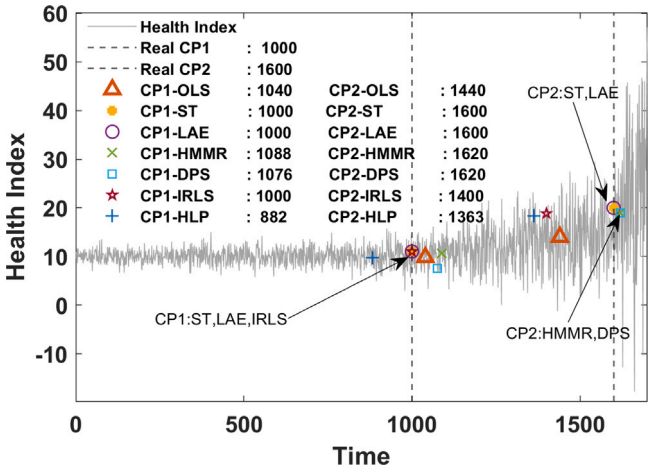


Fig. 5. Changing points detection in the presence of Gaussian noise.

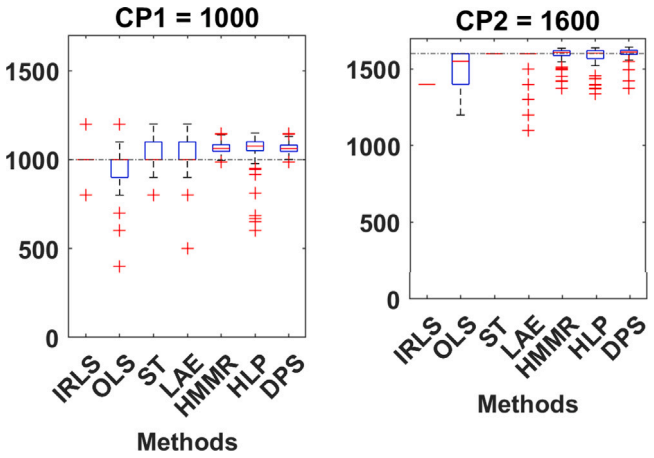


Fig. 6. Monte Carlo analysis for changing point detection in presence of Gaussian noise.

1000. In the case of CP2, most of the methods (excluding IRLS, OLS and ST) tend to have a significant number of outliers on the left side from the true value of 1600.

Also, MSE is calculated based on the differences between true changing point values (CP1 = 1000 and CP2 = 1600) and the estimates from Monte Carlo simulation performed for each method. It is illustrated in Fig. 7. As can be seen in this Fig. 7, IRLS method has the lowest MSE for detecting CP1, while it has the highest MSE for detecting CP2. Also, TS method could detect CP2 very well. However, it has not been able to perform as well as CP2 in detecting CP1. By comparing all the bars, it can be concluded that DPS, HMMR, and ST methods perform the lowest MSE in the case of Gaussian noise.

4.3. Analysis for non-Gaussian noise

In this part, the proposed methodology is applied to the data generated by the proposed model with Student's t distributed random component, considering different values of the parameter ν varying from two marginals: the value closely above 2 which is the limit case below which variance becomes infinite (highly non-Gaussian) and the value large enough so that the distribution is close to Gaussian.

According to the simulation procedure, CP1 and CP2 are equal to Time = 1000 and Time = 1600, respectively. In Fig. 8 we can observe an exemplary simulated trajectory along with the estimated changing points CP1 and CP2 resulting from each of the methods for the case of degrees of freedom $\nu = 2.1$. CP1 is discovered by methods such as ST and LAE precisely, while the rest of the methods could not correctly identify this point; for example, HMMR, DPS, IRLS and HLP discover this point at Time = 347, 348, 600, and 900, respectively; also, OLS detected this point with a delay at Time = 1280. Similarly, ST and LAE precisely detected CP2, and HMMR and DPS discovered this point with a slight delay at Time = 1612, which is acceptable, while IRLS on OLS identified this point at Time = 1400 and 1440, which is very early, and the results of HLP are not valuable.

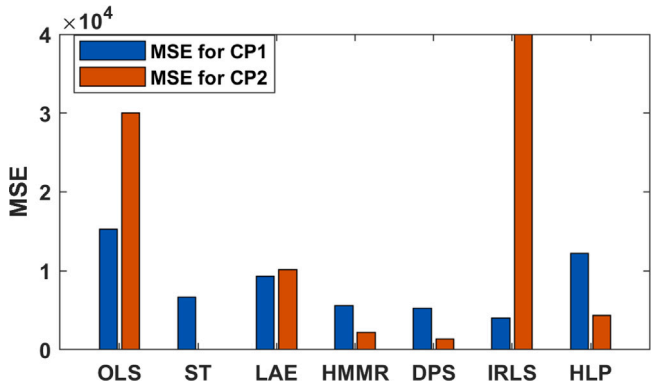


Fig. 7. MSE for Monte Carlo analysis for changing point detection in presence of Gaussian noise.

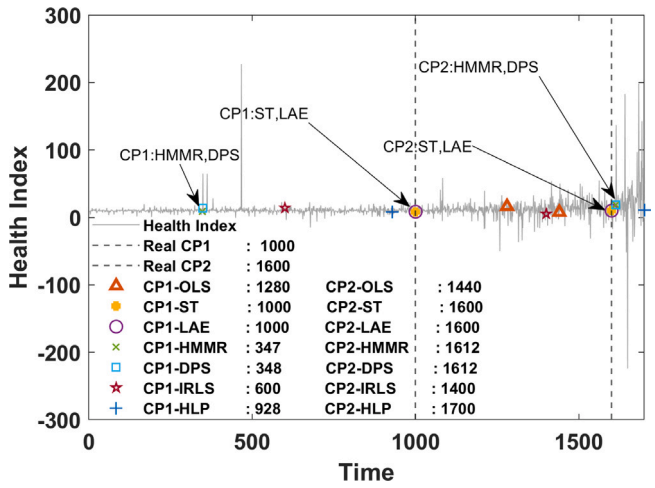


Fig. 8. Segmentation results for non-Gaussian simulation case with $\nu = 2.1$.

To confirm efficacy of the methods, estimation procedure was repeated for 100 simulations from the same model for different levels of non-Gaussian noise. In Fig. 9 the estimation results are visualized using box plots. By analyzing the box plot, as we expect, it can be seen that obviously ST method has the best performance in detecting both changing points(particularly CP2) in presence of different non-Gaussian noise levels.

Also, MSE is calculated based on real changing points (CP1 = 1000 and CP2 = 1600) for all Monte Carlo simulations for different values of ν and shown in Fig. 10. As can be seen in panel (a) of Fig. 10, MSE of many mentioned methods used to detect CP1, such as OLS, HMMR, DPS, and more or less HLP, has increased with decreasing ν (decreasing ν causes increased impulsiveness). However, the reduction of ν value does not have a particular effect on MSE of the robust method such as ST, LAE, and IRLS. Also, it should be mentioned that ST, IRLS, and HLP have the lowest MSE for detecting CP1 among all methods mentioned. Likewise, MSE of methods for CP2 is shown in panel (b) of Fig. 10. MSE of methods such as HMMR, DPS, and more or less OLS increased with decreasing ν, while MSE of methods such as ST, LAE, IRLS, and more or less HLP is not affected by decreasing value of ν. Furthermore, ST method could detect CP2 with the lowest MSE among all methods.

In order to demonstrate universality of the proposed approach, we also verified the methodology for another non-Gaussian distribution, namely the stable distribution [53]. Details of this distribution are presented in Appendix A.1. Random variables from stable distribution (except Gaussian case) have infinite variance; thus, identification of the changing points is much more difficult than in the case considered in this paper, also for the analyzed non-Gaussian distribution. In Appendix A.2 we demonstrate the estimation results of changing points CP1 and CP2 for signals with stable distribution for the selected value of the parameter responsible for non-Gaussian behavior, i.e. α parameter. We recall that $\alpha < 2$ indicates non-Gaussian heavy-tailed behavior of the signal. The results presented in Appendix A.2 clearly indicate the efficiency of the proposed approaches also in this case and confirm the universality of the methodology introduced.

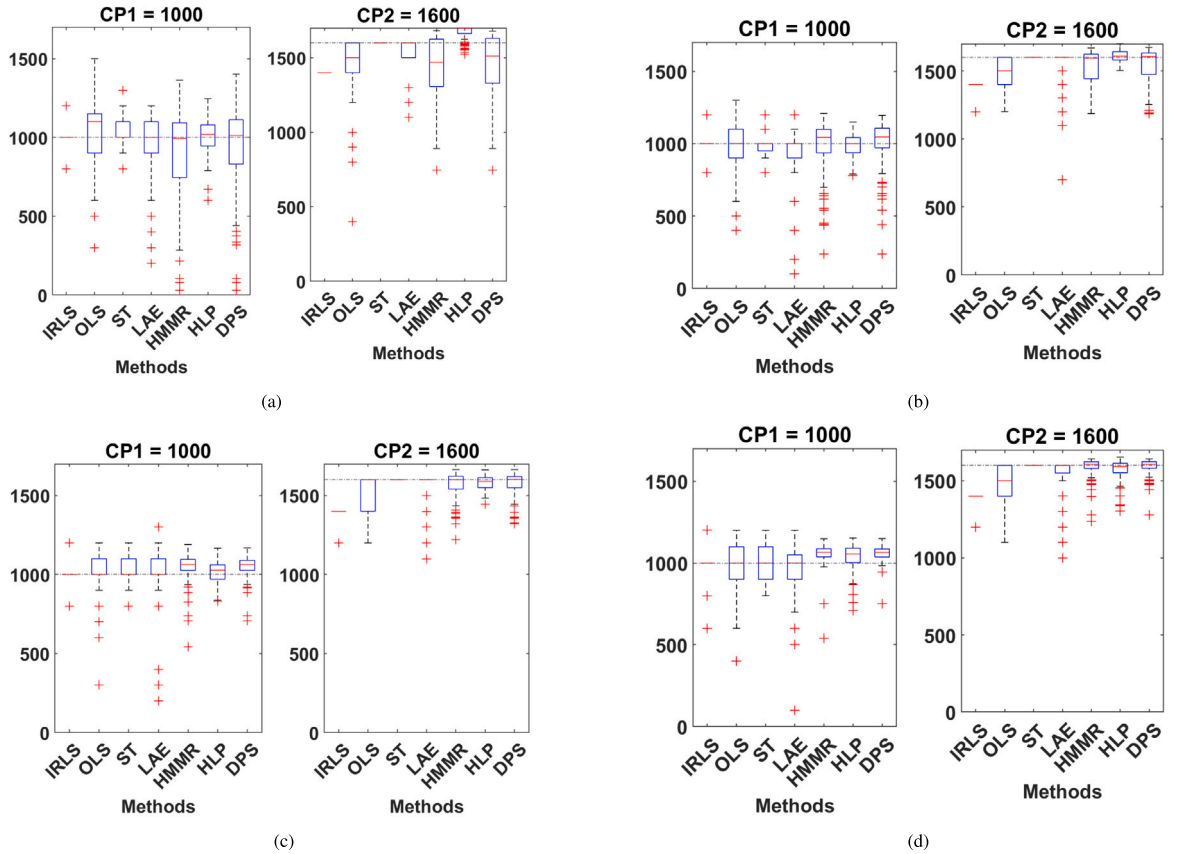


Fig. 9. Monte Carlo analysis for changing point detection in the presence of different levels of non-Gaussian noise, (a) $\nu = 2.1$, (b) $\nu = 3$, (c) $\nu = 5$, (d) $\nu = 10$.

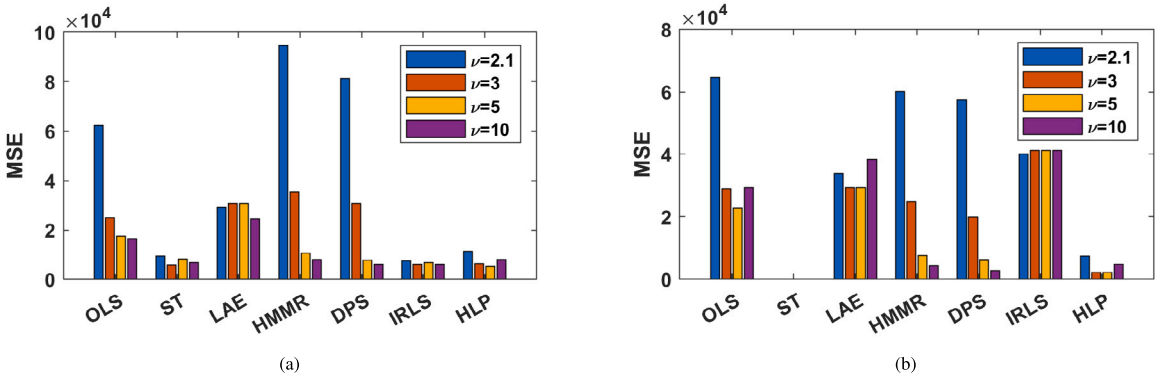


Fig. 10. MSE for Monte Carlo analysis for changing point detection in the presence of different levels of non-Gaussian noise, (a) bar plot of MSE for CP1, (b) bar plot of MSE for CP2.

5. Real data analysis

In this section, we apply and evaluate the proposed methodology for available real data sets. These data are typically employed as benchmark data sets for different papers and competitions and have specific behavior corresponding to noise properties and deterministic trends. In following sections, essential information about objects, experiments, and data will be recalled, and appropriate references are provided.

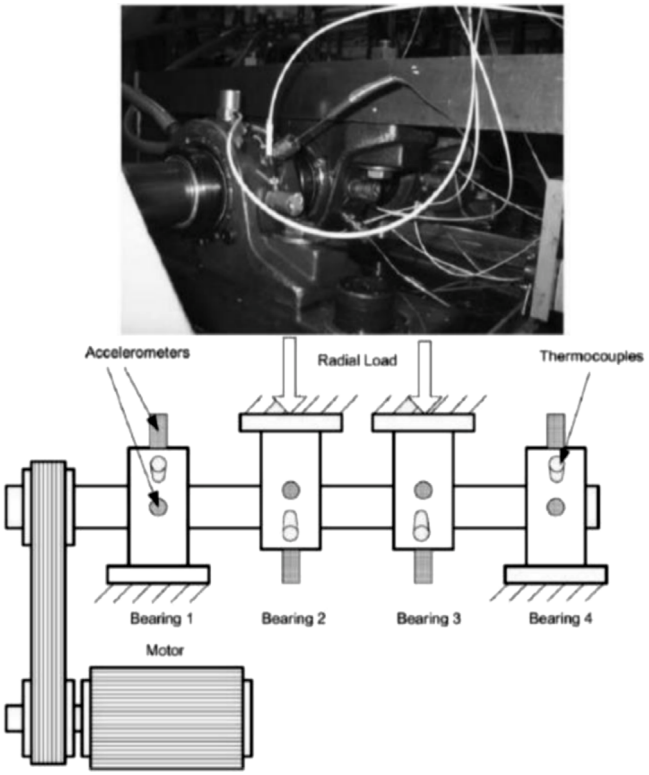


Fig. 11. IMS test rig [61].

5.1. IMS data set

This data set was collected by the Intelligent Maintenance System (IMS) laboratory of the University of Cincinnati. The IMS data set included three subsets of bearing degradation tests. Four Rexnord ZA-2115 double-row bearings were used on a shaft during degradation tests Fig. 11. Accelerometers were mounted on the bearing housings. At the end of the test, degradation patterns are recorded in detail by checking the bearings. Bearings in the IMS data set have a longer and more complicated degradation trend than other benchmark data sets. In some cases, the 'increase-decrease-increase' behavior can also be observed in bearing degradation trends related to the 'self-healing' nature of damage phenomena. Furthermore, this data set has enough frequency resolution of the vibration signals for extracting frequency domain features to monitor bearing degradation process and frequency analysis for condition monitoring. According to data set documentation, each file consists of 20,480 data samples with a sampling frequency of 20 kHz. Furthermore, this data set has been used in many publications for segmentation [54,55], RUL prediction [56–60] and condition monitoring [61].

To apply our proposed methodology to IMS data set, data set number 2 is selected as a case study. This data set was collected from February 12, 2004, 10:32:39 to February 19, 2004, 06:22:39, and includes 984 sets of recorded vibration data. Every of these vibration sets is recorded for 1 s with a 20 kHz sampling rate, and this procedure is repeated every 10 min panel (a) in Fig. 12. In this study, root mean square (RMS) of each vibration set is employed as HI. As shown in panel (b) in Fig. 12 at the beginning of degradation process, vibration amplitude increased due to the impact generated by the initial surface defect, such as cracks or spalling. Then the impact amplitude decreased because the initial surface defect was smoothed by continuous rolling contact. When damage is spread over a broader area, the vibration amplitude increases again.

5.2. FEMTO data set

The FEMTO data set was acquired by Franche-Comté Electronics Mechanics Thermal Science and Optics– Sciences and Technologies institute from a PRONOSTIA platform, Fig. 13. This data set includes 17 historical records that show bearing degradation. Two accelerometers and a temperature sensor are used to acquire acceleration and temperature. Additionally, speed of the shaft was kept stable during the test. In this data set, the bearing failure pattern and the bearing degradation trend are different under the same operating conditions. Furthermore, sampling frequency of the data set is 25 600 Hz, but the length of the signal is 0.1 s, so it is too low for standard frequency analysis and spectrogram. Similarly to the IMS data set, this data set has been

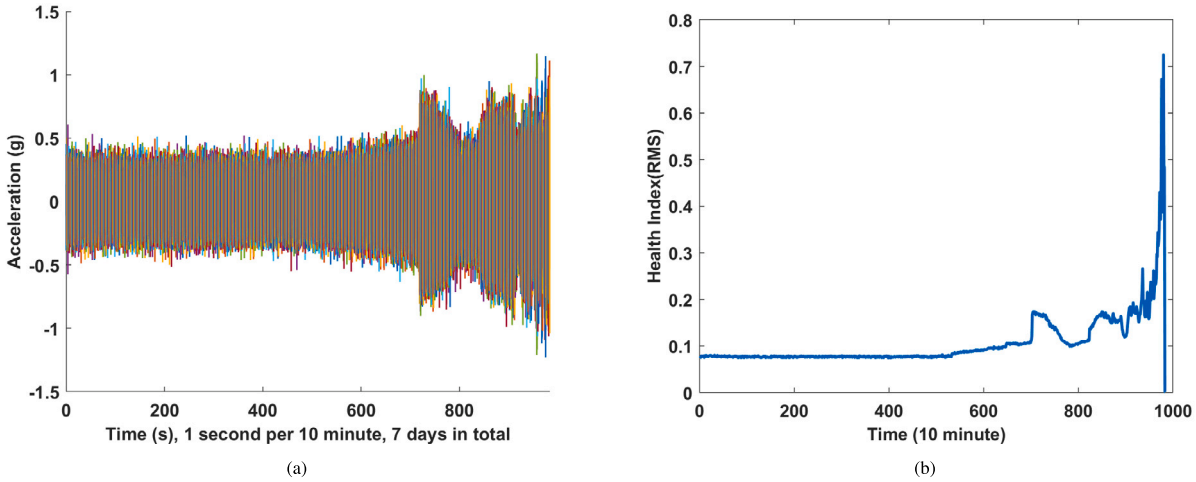


Fig. 12. IMS data set case number 2, (a) raw bearing run-to-failure vibration signals, (b) health index (RMS).

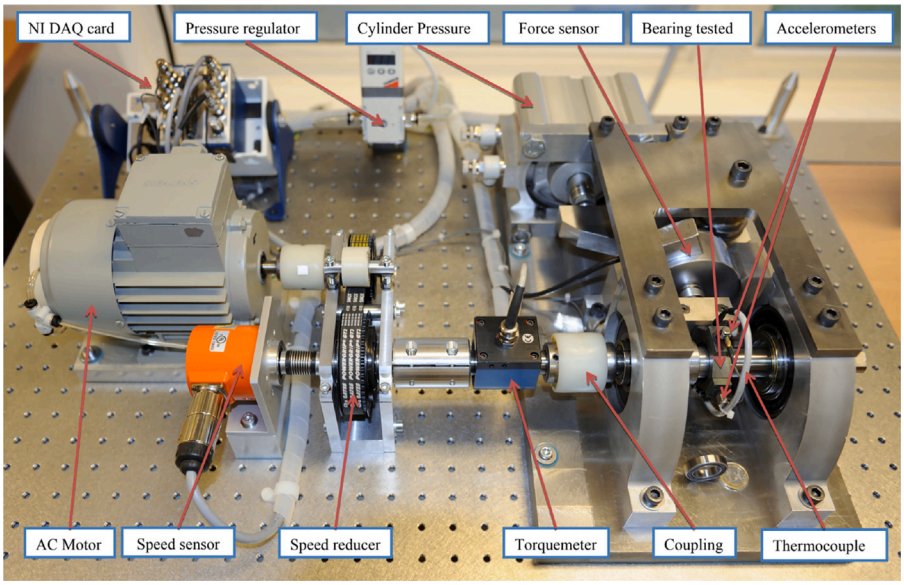


Fig. 13. FEMTO test rig [70].

employed in many publications for health index construction [62–64], segmentation of degradation process [17,65,66] and RUL prediction [67–69].

Bearing1 – 1 is selected as a case study in this research from the FEMTO data set. This data set includes 2803 sets of recorded vibration data. Each of these vibration sets is recorded for 0.1 s with a 25.6 kHz sampling rate (see panel(a) in Fig. 14). This process is repeated every 10 s. The shaft speed is approximately 1800 rpm and the load is equal to 4000 N. Also, in this study, RMS of each vibration set is employed as HI, see panel (b) in Fig. 14.

5.3. Wind turbine data set

This data set is collected from a wind turbine (see Fig. 15). The sensor has been mounted on a high-speed bearing shaft of the 2.2 MW power wind turbine. This data set was acquired over more than 50 days using two different ways. In the first approach, raw vibration signal is acquired for 6 s with 100000 Hz sampling frequency per day for +50 days (see panel (a) in Fig. 16) and for the second version, the bearing inner race energy is calculated every 10 min for +50 days (see panel (b) in Fig. 16). For more information on the methodology used to calculate the inner race energy, see [71].

In the end, the inner race fault has occurred, which has been proven by inspection, see Fig. 15. Bearing type used during the test is $32222 - J2$- SKF. It should be mentioned that this data set has been used for prognosis by several papers [71–73].

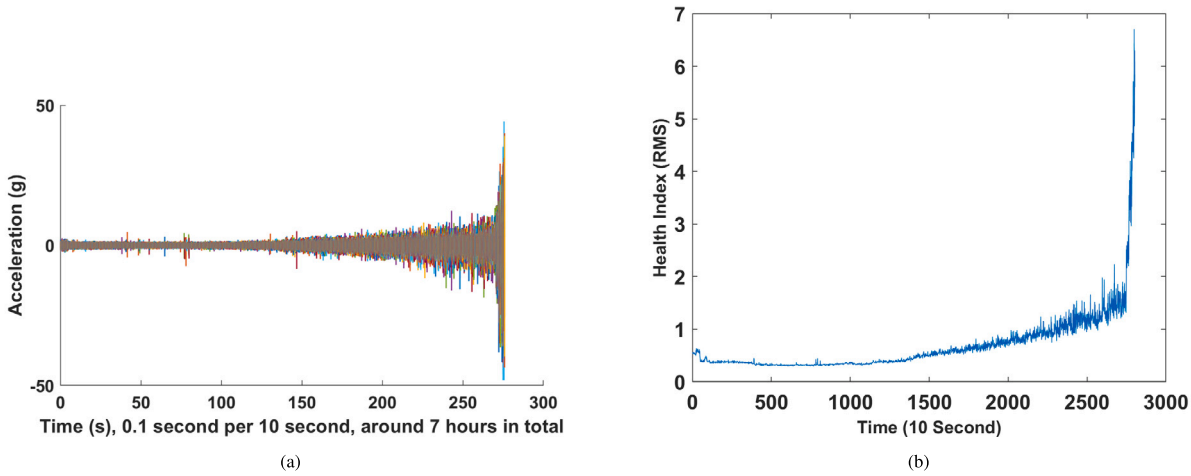


Fig. 14. FEMTO data set case number 1, (a) Raw bearing run-to-failure vibration signals, (b) health index (RMS).

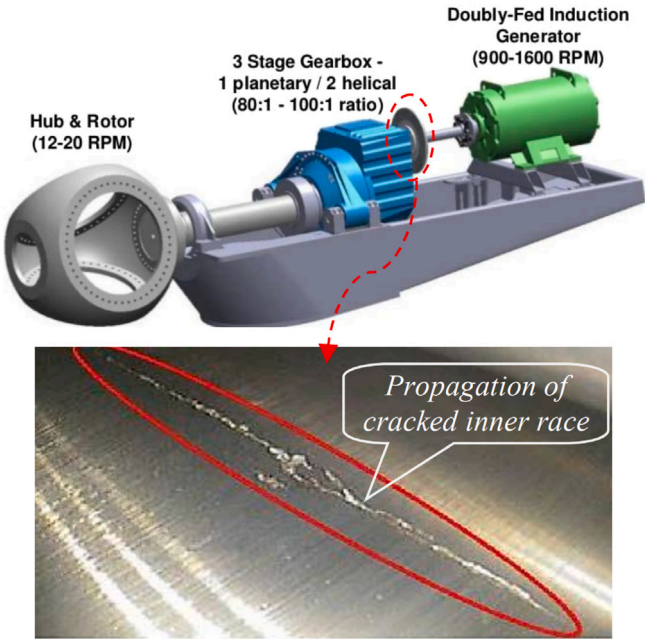


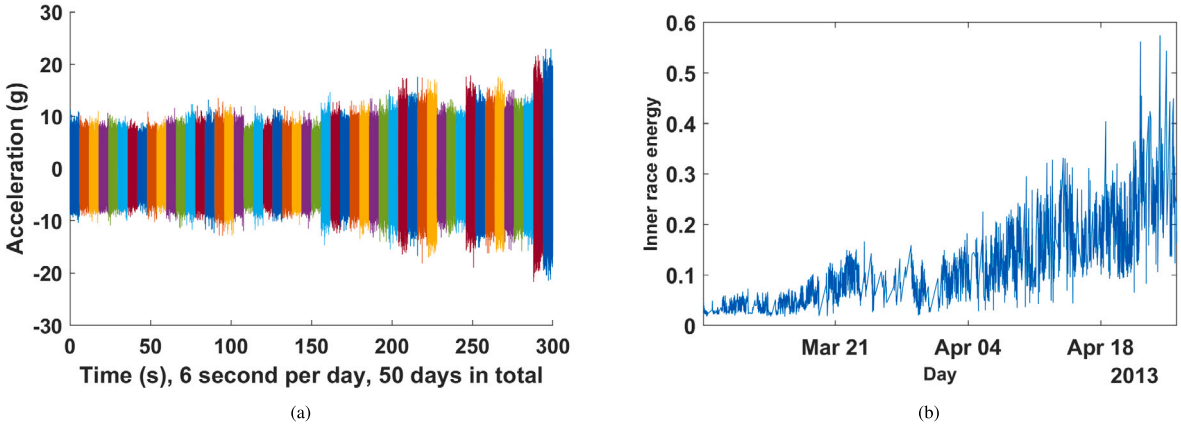
Fig. 15. Wind turbine test rigs [71].

5.4. Results and discussion

In this subsection, we presented the results of our proposed methodology for three real data sets (IMS, FEMTO, and wind turbine data set). We apply our robust approach to every data set and compare the results with conventional methods.

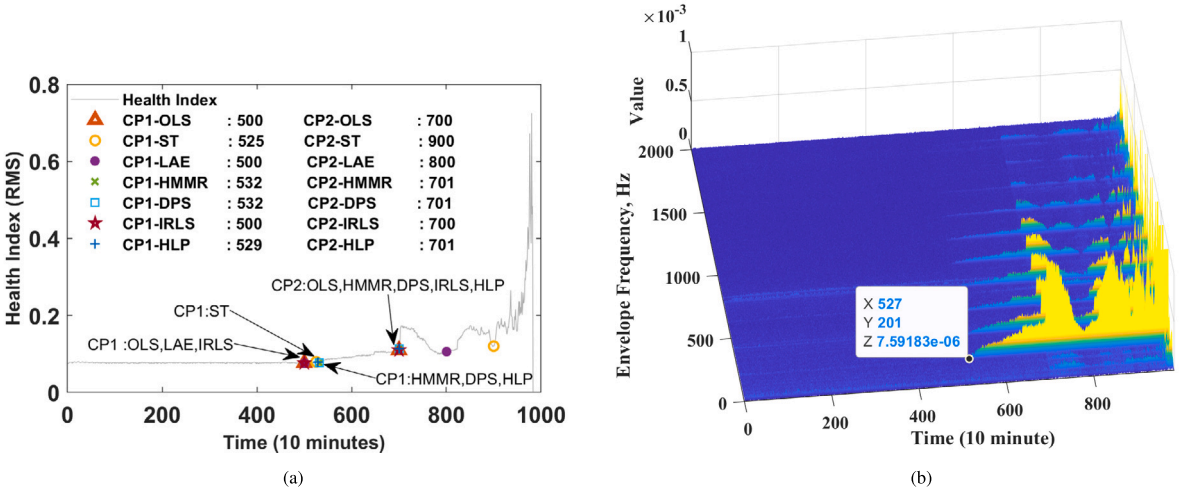
5.4.1. Results for IMS data set

Results of the segmentation methods for this case study are presented in panel (a) of Fig. 17. As can be seen in panel (a) of Fig. 17, most methods detected first changing point (CP1) (boundary point between healthy stage and degradation stage) between Time = 500 and 560. By visual check of HI and the results, it looks like most of the methods gave adequate results. For the second changing point (boundary point between degradation stage and critical stage trend), many methods such as OLS, HMMR, DPS, IRLS and HLP detected this point around Time = 700, where indeed HI presents a large increase of value. The same situation is for the LAE method — it has identified Time = 800 and indeed after that point HI values are growing rapidly. Unfortunately, both mentioned cases seem to be not correct/true CP2 points in a global sense, as HI values started decreasing. This is the reason that we



(a) (b)

Fig. 16. Wind turbine data set, (a) raw bearing run-to-failure vibration signals, (b) health index (Inner race energy).



(a) (b)

Fig. 17. Changing points detection IMS data set case number 2, (a) result of changing points detection method, (b) envelope spectrum of the raw vibration signal.

claim CP2 = 900 identified by the ST method is correct as the trend of the data after that point suggests exponential degradation for a wide time span. Also, to confirm our results, envelope analysis is applied for each vibration set, and the result is sequentially plotted in panel (b) of Fig. 17. For more details about envelope analysis, we refer to [74–77].

As can be seen in panel (b) of Fig. 17 after Time = 527 harmonic frequency has appeared – namely the Ball Pass Frequency Outer (BPFO) fundamental harmonic – which means the bearing went to the degradation stage. By comparison of the envelope analysis and segmentation methods, it can be concluded that ST and HLP methods provide the best result.

5.4.2. Results for FEMTO data set

Results of the segmentation methods for this case study are presented in Fig. 18. As can be seen, this data set perfectly follows the idea of 3 stages: for Time = 0 to c.a. 1300 it is nearly flat, then up to Time = 2700 there is a linear increase and then rapid growth occurs. Some noise, especially in the middle stage, is seen, and some minor outliers can also be noticed. Thus, we consider it to be a trend with nearly Gaussian noise. As shown in Fig. 18 most of the algorithms detected first changing point (boundary point between healthy and degradation stage) between Time = 1200 and 1400, except HMM-based algorithms such as HMMR and HLP that discovered first changing point much later than it occurs in reality. Based on the visual check first changing point should be located around Time = 1300. That means that ST method could detect this point with acceptable error. It should be considered that transferring between healthy and degradation stages is smooth, and it is not easy to detect this point. According to Fig. 18 second changing point should be somewhere around Time = 2700, where the health index is starting to dramatically increase — in global sense. Therefore, it can be concluded that ST, OLS, HMMR, and HLP methods could detect this point as well. However, LAE, DPS, and IRLS methods have not been able to discover this point properly.

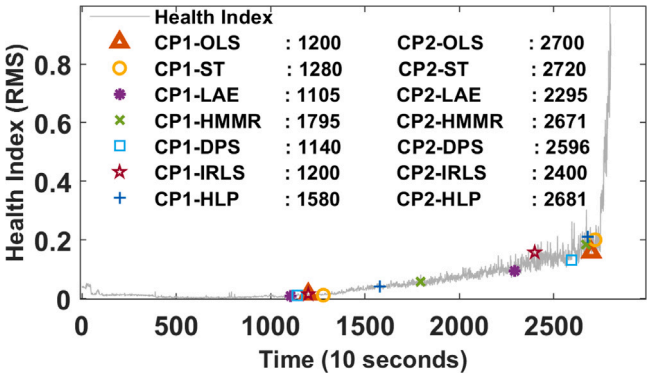


Fig. 18. Result of changing points detection methods for the FEMTO data set.

5.4.3. Results for wind turbine data set

To apply the proposed approach to the wind turbine data set, inner race energy is selected as a health index. As can be seen in panel (b) of Fig. 16 this data set is include a lot of outliers. Thus, we consider it as a trend with non-Gaussian noise. In addition, a few fluctuations have appeared on HI may that have corresponded to variation of the load or some phenomena like self-healing. This specific behavior in real cases makes a significant challenge to segmentation and prediction of RUL. In Fig. 19 results of the segmentation methods for this case study are presented. CP1 (changing point between healthy stage and degradation stage) is detected by all methods somewhere between 15 March and 21 March; for instance, OLS, ST, LAE and IRLS discovered this point on 15 March, while HMMR, DPS, and HLP detected this point on 18 March, 18 March, 19 March, respectively.

Looking at Fig. 19, 15 March seems a more logical candidate for CP1 because, after this day, HI is dramatically growing until 21 April, which means that this point is discovered by HMMR, HLP and DPS with a delay. CP2 (boundary point between degradation stage and critical stage trend) for HMMR, DPS, and HLP was detected on 8 April. Unfortunately, all mentioned methods seem to detect wrong CP2 points in a global sense. It is also wrong for LAE which detected CP2 when HI values started decreasing. Meanwhile, ST and IRLS discovered it on 20 April. It can be concluded that CP2 = 20 April is the correct CP2 as the trend of the data after that point suggests exponential growth with less fluctuation for a wide time span. Another version of this data set is used to validate our results. As discussed before, this version of the data set is made up of a raw vibration signal recorded for six seconds every day with a sampling rate of 100k; see panel (a) in Fig. 20. So it means that it has excellent potential to apply various frequency analyses to do condition monitoring. Envelope analysis is applied for each vibration set, and the result is sequentially plotted in panel (b) of Fig. 20. As can be seen in panel (b) of Fig. 20 after 14 March, harmonic frequency – namely the Ball Pass Frequency Inner (BPFI) fundamental harmonic – has appeared, which means the bearing went to degradation stage. Also, after 22 April, the harmonic frequency amplitude increased dramatically, which can be considered CP2. By comparing results of the segmentation method, it can be found that IRLS and ST methods have the best results for this case.

5.4.4. Discussion

The percentage error for detecting changing points in all actual data sets is demonstrated in Fig. 21. As we can see in both subfigures for both CP1 and CP2, the ST method has the lowest percentage error. In contrast, error percentage for the rest of the methods has fluctuated on different data sets, which can include the fact that considering non-Gaussian noise distribution improves the process of detecting changing points in long-term condition monitoring data. Moreover, Appendix B presents a comprehensive table detailing computational costs associated with each methodology when applied to real-world datasets. This information provides valuable insights into computational requirements of the different approaches under consideration.

It should be noted that in this paper, particularly in real data analysis, we tried to present envelope analysis to give this option to readers to check what exactly happens in the frequency domain at given time and around the points that are detected by different methods (how much can they provide us a physical meaning). Also, we confirm that condition monitoring based on non-Gaussian nature of the data acquired from the machine may need more advanced signal processing techniques [33,35,78] instead of basic envelope analysis.

Also, we should clarify that one objective of this study is to provide a comprehensive analysis of full datasets in order to deepen our understanding of degradation process and develop expertise in this domain, rather than focusing on evaluation of online methods, which probably will be much harder to achieve. However, it does not mean that this work is useless. In many cases, one may deal with many machines of the same type. We can use historical data sets to build knowledge about machine behavior. It is clear that selecting such CPx points could be done by experts but it is non-objective and prone to bias. In this paper, several methods are tested and in easy cases results provided by various methods are similar. For more complicated cases – as this from Fig. 17(a) – the role of the expert is limited to selecting the best option from 3 suggestions instead of analysis of the whole data set. In addition, it is essential to emphasize the significance of evaluating post-factum segmentation methods for tasks such as model identification and labeling. Results obtained from these evaluations can serve as valuable training data for online methods, which can employ a supervised learning approach and operate in real-time. It is important to note that the segmentation task becomes even more

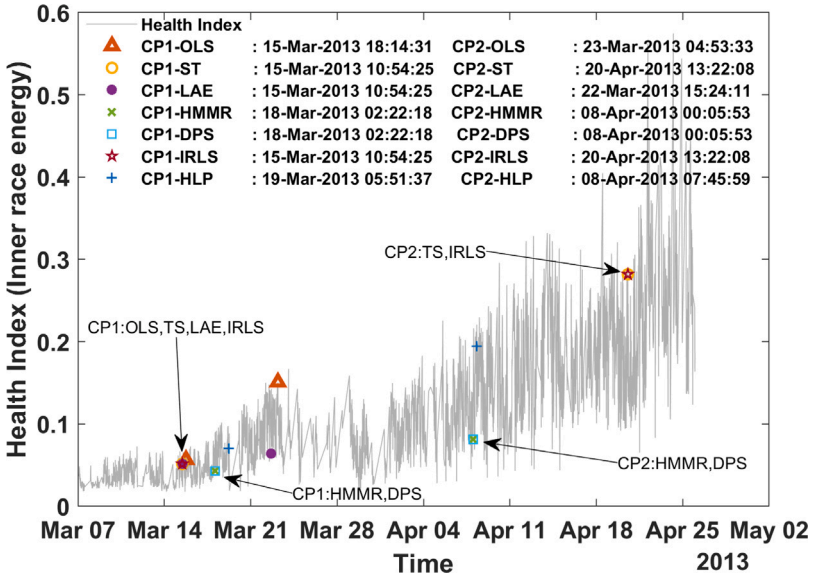
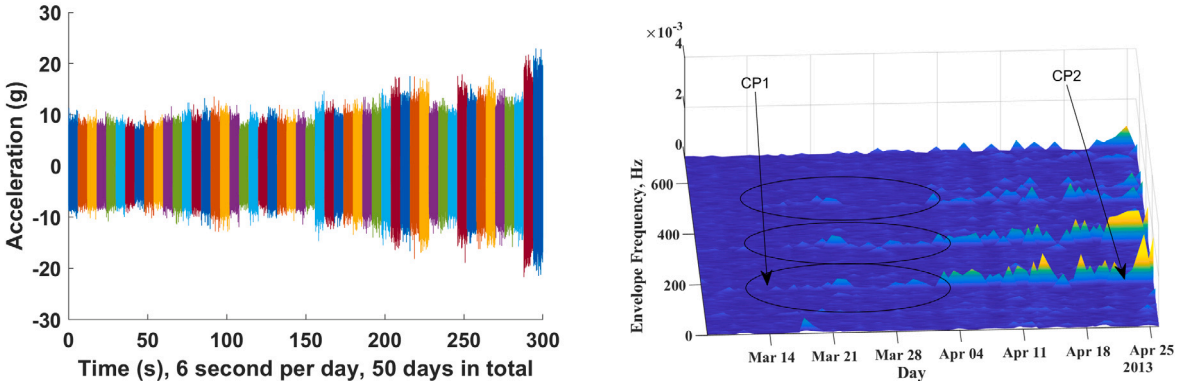
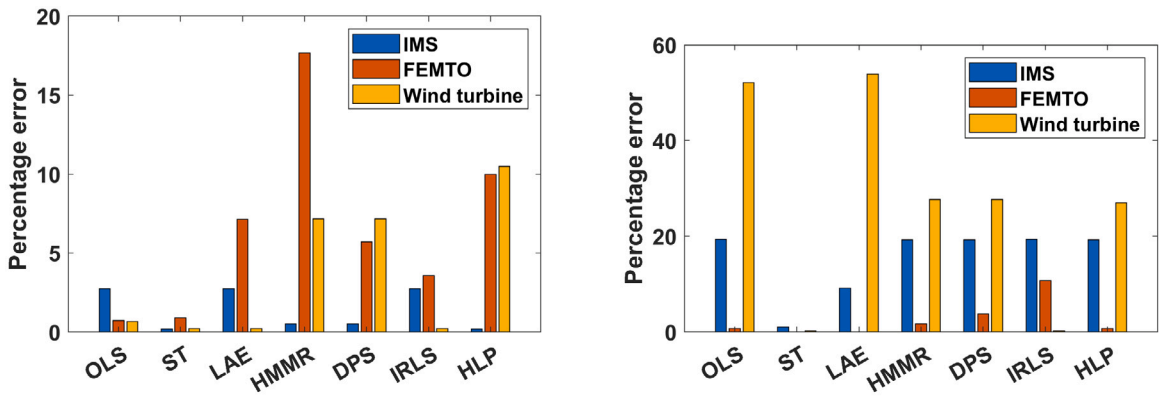


Fig. 19. Changing points detection of wind turbine data set.



(a) (b)

Fig. 20. Changing points detection wind turbine, (a) raw vibration data, (b) envelope spectrum map.



(a) (b)

Fig. 21. Percentage error for real data sets changing points detection, (a) percentage error for CP1, (b) percentage error for CP2.

challenging due to the presence of non-Gaussian noise and time-varying characteristics in the data. By addressing these challenges, we believe this paper establishes a crucial foundation for future research in the development of effective online methods for real-time diagnostics. Conclusively, it is crucial to underscore that the three-regime model represents an idealized scenario, whereas real-world machinery might not precisely adhere to this model. Various other influences, such as those encountered in the IMS and wind turbine datasets, may introduce deviations in the curve's shape. The emergence of a non-linear pattern within the health index presents an additional challenge alongside the influence of non-Gaussian noise. It is our view that the field of PHM would benefit from paying more attention to this aspect in future endeavors. Specifically, a greater focus on constructing a health index characterized by more consistently monotonic trends could mitigate the impact of non-linearity.

6. Conclusions

In the paper, a segmentation problem of long-term degradation (health index) data from condition monitoring systems has been investigated. The long-term data describes degradation process. There are several approaches that assume a model of such degradation. Here, we follow the idea of 3 stages with different properties (see Fig. 3). In practical application, it is a really crucial task to detect when a machine is changing its condition from a healthy stage to a degradation stage (warning) and degradation stage to a critical stage (alarm), and it is the basis for further analysis such as prognostics. Also, in many papers, researchers focus on the last segment, but they do not mention how they have found the division between stages 2 and 3 and whether this detected point coincides with the true start time of the rapid development of the fault in the machine. Therefore, in this paper, we tested various techniques for CPx finding, as CPx is a basis for identifying data segments with different properties. The achievements of this paper can be summarized as follow:

- A model of HI data was proposed as a three segments sequence with non-Gaussian noise (Student's t distributed) to describe degradation process, which can be used to simulate the artificial data set.
- Moreover, we showed that non-Gaussianity of the model makes detection efficiency worse.
- We applied selected methods to simulate and analyze signals with Gaussian and non-Gaussian noise. We used Monte Carlo simulation and box plot-based visualization to provide meaningful results. In the case of Gaussian noise, most of the mentioned methods almost have been able to detect changing points, although a few methods, such as HLP and ST, had the lowest error. However, the situation is entirely different in the presence of non-Gaussian noise. As expected, the efficacy of methods derived on the basis of Gaussian distribution decreased with the increase in impulsiveness of HI. In contrast, the efficacy of robust methods such as ST, LAE, and IRLS have not been significantly influenced by increasing impulsiveness.
- We have noticed that detection of CP1 is much more difficult than that of CP2. Statistical properties evolve in time over whole life of the machine. The difference between the end of stage 1 and the beginning of stage 2 is not so clear as between stage 2 and stage 3, as we change trend from linear to exponential.

We applied the procedure to real data sets known in the community as benchmark (reference) data sets (namely IMS, FEMTO, and data from wind turbine drive). Additionally, we presented a waterfall plot of the envelope spectrum calculated for available raw vibration data for IMS and wind turbine data sets. Such a 3D plot demonstrates how the envelope spectrum changes allow validating visually that detected changing points are correct. Also, the results obtained from applying the mentioned methods to real data sets and comparing them with the results of envelope analysis show that the ST method has the most efficiency in detecting changing points. Furthermore, because the investigated data sets are benchmark data sets, many papers use them as references every year. Researchers can use the results of the real data set section, particularly 3D plots to compare research results.

Also, it is noteworthy that the ST method, developed based on non-Gaussian distribution, had the proper efficiency in detecting changing points in both simulated and real data sets. This can be a point to consider for a researcher working in the field of long-term condition monitoring data analysis in a harsh environment to choose noise distribution. Although there are other noise distributions to describe impulsive behavior of degradation process, which may be more compatible with this procedure, this issue deserves more research in the future.

Segmentation has a crucial role in condition-based maintenance and it can be considered as a pre-processing step of larger analytical pipeline. With convincing results of segmentation, one can start describing/modeling the data corresponding to given segments and building knowledge and expertise about the process and/or the machine. Finally, based on the model one can make a prediction which is crucial for RUL analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors (Hamid Shiri) gratefully acknowledge the European Commission for its support of the Marie Sklodowska Curie program through the ETN MOIRA project (GA 955681).

Project no. POIR.01.01.01-00-0350/21 entitled "A universal diagnostic and prognostic module for condition monitoring systems of complex mechanical structures operating in the presence of non-Gaussian disturbances and variable operating conditions" co-financed by the European Union from the European Regional Development Fund under the Intelligent Development Program. The project is carried out as part of the competition of the National Center for Research and Development no: 1/1.1.1/2021 (Szybka Ścieżka) - Agnieszka Wylomanska and Radoslaw Zimroz

Appendix A. Result for the signals with stable noise

A.1. Stable distribution

In non-Gaussian case, symmetric stable distribution have been selected as an example of heavy-tailed distributions [53].

In general, the stable distribution is defined by its characteristic function and is characterized by four parameters: α (stability), β (skewness), σ (scale), and μ (location). However, for the symmetric case with a standardized scale, we assume $\beta = \mu = 0$, $c = 1$, and the corresponding characteristic function is given by the formula:

$$E[e^{itX}] = e^{-|t|^{\alpha}}. \tag{A.1}$$

The parameter α is called the stability index and takes the value from $(0, 2]$ interval. It should be noted that the stable distribution reduces to Gaussian distribution when $\alpha = 2$. In case when α decreases, the distribution becomes significantly non-Gaussian and heavy-tailed, [79]. In the simulation study, we assume that the stability index is $\alpha = 1.9$.

A.2. Monte Carlo analysis for stable distributed signals

The estimation procedure was repeated for 100 simulations of the same model (similar as in Section 4.1). However in this case noise $\tilde{R}(t)$ was generated from stable distribution with $\alpha = 1.9$, instead of Student's t distribution. In Fig. A.22, the estimation results are visualized using two box plots. The vertical lines of each blue box correspond to IQR (the interval between Q1 and Q3). In addition, outliers can be observed. Based on the CP1 box plot, the median value of the estimator obtained by IRLS, OLS, ST and LAE is at CP1 = 1000, which means that these methods can accurately detect first changing point in most cases. In the case of CP2 box plot, the median value of the estimator is close to CP2 = 1600 for most methods, except IRLS and OLS. As we can see, the robust techniques (ST and LAE) are also efficient for both changing points in the case when the Student's t distribution of noise is replaced with stable distribution. This assures us about their potential effectiveness in general non-Gaussian case.

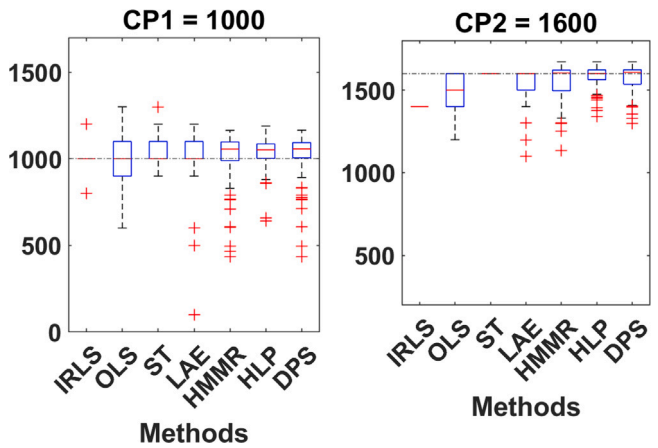


Fig. A.22. Monte Carlo analysis for changing point detection in the presence of non-Gaussian noise (stable noise with $\alpha = 1.9$).

Table B.2
Computational cost.

	OLS	ST	LAE	HMMR	DPS	IRLS	HLP
FEMTO data set	1.96 (s)	2.89 (s)	1.98 (s)	1.11 (s)	118.09 (s)	16.85 (s)	0.38 (s)
Wind turbine data set	6.97 (s)	8.91 (s)	7.13 (s)	1.17 (s)	399.07 (s)	40.02 (s)	0.68 (s)
IMS data set	0.13 (s)	0.34 (s)	0.27 (s)	0.47 (s)	6.78 (s)	2.65 (s)	0.34 (s)

Appendix B. Computational cost

Table B.2 demonstrates the computational cost of all methods for real data sets. The proposed method is implemented with Matlab 2021b, and the hardware property of the system for this implementation are as follows: Processor: Intel (R) Core (TM) i7-10750H CPU @ 2.60 GHz 2.59 GHz and Ram: 32.0 GB.

As anticipated, the methodologies employing dynamic programming, such as DPS, and iterative methods, such as IRLS, exhibit the highest computational costs, respectively. In contrast, the methodologies based on Hidden Markov Models (HMM), such as HHMR and HLP, demonstrate the lowest computational costs among all the approaches. Our analysis also highlights the enhanced accuracy of the Student's t (ST) method, which, despite its superior performance, maintains a moderate range of computational costs.

References

[1] B.A. Weiss, J. Pellegrino, M. Justiniano, A. Raghunathan, Measurement Science Roadmap for Prognostics and Health Management for Smart Manufacturing Systems, National Institute of Standards and Technology, 2016, pp. 100–102.
[2] F. Moosavi, H. Shiri, J. Wodecki, A. Wyłomańska, R. Zimroz, Application of machine learning tools for long-term diagnostic feature data segmentation, Appl. Sci. 12 (13) (2022) 6766.
[3] S. Aminikhanghahi, D.J. Cook, A survey of methods for time series change point detection, Knowl. Inf. Syst. 51 (2) (2017) 339–367.
[4] M. Mudelsee, Trend analysis of climate time series: A review of methods, Earth-Sci. Rev. 190 (2019) 310–322.
[5] B. Zhou, H. Ye, A study of polynomial fit-based methods for qualitative trend analysis, J. Process Control 37 (2016) 21–33.
[6] A. Samé, F. Chamroukhi, G. Govaert, P. Aknin, Model-based clustering and segmentation of time series with changes in regime, Adv. Data Anal. Classif. 5 (4) (2011) 301–321.
[7] S. Lee, Y. Jeong, D. Park, B.-J. Yun, K.H. Park, Efficient fiducial point detection of ECG QRS complex based on polygonal approximation, Sensors 18 (12) (2018) 4502.
[8] A. Shirvani, K. Arpe, M. Jahandideh, Analysis of trends and change points in meteorological variables over the south of the Caspian Sea, Theor. Appl. Climatol. 141 (3) (2020) 959–966.
[9] J.-Y. Cheung, G. Stephanopoulos, Representation of process trends—Part I. A formal representation framework, Comput. Chem. Eng. 14 (4–5) (1990) 495–510.
[10] K. Gąsior, H. Urbańska, A. Grzesiek, R. Zimroz, A. Wyłomańska, Identification, decomposition and segmentation of impulsive vibration signals with deterministic components—a sieving screen case study, Sensors (Switzerland) 20 (19) (2020) 1–20.
[11] D. Kucharczyk, A. Wyłomańska, J. Obuchowski, R. Zimroz, M. Madziarz, Stochastic modelling as a tool for seismic signals segmentation, Shock Vib. 2016 (2016).
[12] A. Grzesiek, K. Gasior, A. Wyłomańska, R. Zimroz, Divergence-based segmentation algorithm for heavy-tailed acoustic signals with time-varying characteristics, Sensors 21 (24) (2021).
[13] A. Alkaya, İ. Eker, Variance sensitive adaptive threshold-based PCA method for fault detection with experimental application, ISA Trans. 50 (2) (2011) 287–302.
[14] O. Fink, E. Zio, U. Weidmann, A classification framework for predicting components' remaining useful life based on discrete-event diagnostic data, IEEE Trans. Reliab. 64 (3) (2015) 1049–1056.
[15] M. Schlechtingen, I.F. Santos, Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection, Mech. Syst. Signal Process. 25 (5) (2011) 1849–1875.
[16] L. Hu, N. Hu, X. Zhang, F. Gu, M. Gao, Novelty detection methods for online health monitoring and post data analysis of turbopumps, J. Mech. Sci. Technol. 27 (7) (2013) 1933–1942.
[17] J.K. Kimotho, C. Sondermann-Wölke, T. Meyer, W. Sextro, Machinery prognostic method based on multi-class support vector machines and hybrid differential evolution–particle swarm optimization, Chem. Eng. Trans. 33 (2013).
[18] E. Sutrisno, H. Oh, A.S.S. Vasan, M. Pecht, Estimation of remaining useful life of ball bearings using data driven methodologies, in: 2012 IEEE Conference on Prognostics and Health Management, IEEE, 2012, pp. 1–7.
[19] Y. Hu, H. Li, X. Liao, E. Song, H. Liu, Z. Chen, A probability evaluation method of early deterioration condition for the critical components of wind turbine generator systems, Mech. Syst. Signal Process. 76 (2016) 729–741.
[20] P. Tamilselvan, Y. Wang, P. Wang, Deep belief network based state classification for structural health diagnosis, in: 2012 IEEE Aerospace Conference, IEEE, 2012, pp. 1–11.
[21] J. Liu, F. Lei, C. Pan, D. Hu, H. Zuo, Prediction of remaining useful life of multi-stage aero-engine based on clustering and LSTM fusion, Reliab. Eng. Syst. Saf. 214 (2021) 107807.
[22] J. Singh, A. Darpe, S.P. Singh, Bearing remaining useful life estimation using an adaptive data-driven model based on health state change point identification and K-means clustering, Meas. Sci. Technol. 31 (8) (2020) 085601.
[23] W. Mao, J. He, B. Sun, L. Wang, Prediction of bearings remaining useful life across working conditions based on transfer learning and time series clustering, IEEE Access 9 (2021) 135285–135303.
[24] A. Giantomassi, F. Ferracuti, A. Benini, G. Ippoliti, S. Longhi, A. Petrucci, Hidden Markov model for health estimation and prognosis of turbofan engines, in: International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Vol. 54808, 2011, pp. 681–689.
[25] F. Sloukia, M. El Aroussi, H. Medromi, M. Wahbi, Bearings prognostic using mixture of gaussians hidden markov model and support vector machine, in: 2013 ACS International Conference on Computer Systems and Applications (AICCSA), IEEE, 2013, pp. 1–4.

[26] J. Janczura, T. Barszcz, R. Zimroz, A. Wyłomańska, Machine condition change detection based on data segmentation using a three-regime, α-stable Hidden Markov Model, Measurement (2023) 113399.
[27] C.K.R. Lim, D. Mba, Switching Kalman filter for failure prognostic, Mech. Syst. Signal Process. 52 (2015) 426–435.
[28] H. Shiri, P. Zimroz, J. Wodecki, A. Wyłomańska, R. Zimroz, K. Szabat, Using long-term condition monitoring data with non-Gaussian noise for online diagnostics, Mech. Syst. Signal Process. 200 (2023) 110472.
[29] L. Cui, X. Wang, Y. Xu, H. Jiang, J. Zhou, A novel switching unscented Kalman filter method for remaining useful life prediction of rolling bearing, Measurement 135 (2019) 678–684.
[30] K. Gong, X. Chen, Influence of non-Gaussian wind characteristics on wind turbine extreme response, Eng. Struct. 59 (2014) 727–744.
[31] K.R. Gurley, M.A. Tognarelli, A. Kareem, Analysis and simulation tools for wind engineering, Probab. Eng. Mech. 12 (1) (1997) 9–31.
[32] A. Kareem, J. Zhao, Analysis of non-Gaussian surge response of tension leg platforms under wind loads, 1994.
[33] J. Hebda-Sobkowicz, R. Zimroz, A. Wyłomańska, J. Antoni, Infogram performance analysis and its enhancement for bearings diagnostics in presence of non-Gaussian noise, Mech. Syst. Signal Process. 170 (2022) 108764.
[34] J. Nowicki, J. Hebda-Sobkowicz, R. Zimroz, A. Wyłomańska, Dependency measures for the diagnosis of local faults in application to the heavy-tailed vibration signal, Appl. Acoust. 178 (2021) 107974.
[35] J. Wodecki, A. Michalak, R. Zimroz, Local damage detection based on vibration data analysis in the presence of Gaussian and heavy-tailed impulsive noise, Measurement 169 (2021) 108400.
[36] J. Hebda-Sobkowicz, J. Nowicki, R. Zimroz, A. Wyłomańska, Alternative measures of dependence for cyclic behaviour identification in the signal with impulsive noise—Application to the local damage detection, Electronics 10 (15) (2021) 1863.
[37] P. Kruczek, R. Zimroz, A. Wyłomańska, How to detect the cyclostationarity in heavy-tailed distributed signals, Signal Process. 172 (2020) 107514.
[38] W. Zulawinski, K. Maraj-Zygmat, H. Shiri, A. Wylomanska, R. Zimroz, Framework for stochastic modelling of long-term non-homogeneous data with non-Gaussian characteristics for machine condition prognosis, Mech. Syst. Signal Process. 184 (2023) 109677.
[39] D. Zhang, An adaptive procedure for tool life prediction in face milling, Proc. Inst. Mech. Eng. J 225 (11) (2011) 1130–1136.
[40] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, J. Lin, Machinery health prognostics: A systematic review from data acquisition to RUL prediction, Mech. Syst. Signal Process. 104 (2018) 799–834.
[41] L. Zhang, Z. Mu, C. Sun, Remaining useful life prediction for lithium-ion batteries based on exponential model and particle filter, IEEE Access 6 (2018) 17729–17740.
[42] C. Pan, Y. Chen, L. Wang, Z. He, Lithium-ion battery remaining useful life prediction based on exponential smoothing and particle filter, Int. J. Electrochem. Sci. 14 (9) (2019) 9537–9551.
[43] C. Lim, D. Mba, Switching Kalman filter for failure prognostic, Mech. Syst. Signal Process. 52–53 (1) (2015) 426–435.
[44] L. Cui, X. Wang, H. Wang, J. Ma, Research on remaining useful life prediction of rolling element bearings based on time-varying Kalman filter, IEEE Trans. Instrum. Meas. 69 (6) (2019) 2858–2867.
[45] L. Reuben, D. Mba, Diagnostics and prognostics using switching Kalman filters, Struct. Health Monit. 13 (3) (2014) 296–306.
[46] L.C.K. Reuben, D. Mba, Diagnostics and prognostics using switching Kalman filters, Struct. Health Monit. 13 (3) (2014) 296–306.
[47] F. Chamroukhi, A. Samé, G. Govaert, P. Aknin, Time series modeling by a regression approach based on a latent process, Neural Netw. 22 (5–6) (2009) 593–602.
[48] F. Chamroukhi, A. Samé, G. Govaert, P. Aknin, A regression model with a hidden logistic process for feature extraction from time series, in: 2009 International Joint Conference on Neural Networks, IEEE, 2009, pp. 489–496.
[49] P.J. Huber, Robust statistics, in: International Encyclopedia of Statistical Science, Springer, 2011, pp. 1248–1251.
[50] J.O. Street, R.J. Carroll, D. Ruppert, A note on computing robust regression estimates via iteratively reweighted least squares, Amer. Statist. 42 (2) (1988) 152–154.
[51] Least absolute deviation estimates in autoregression with infinite variance, 16.
[52] R. Li, S. Nadarajah, A review of Student's t distribution and its generalizations, Empir. Econ. 58 (3) (2020) 1461–1490.
[53] G. Samorodnitsky, M. Taqqu, Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance, Chapman and Hall, 1994.
[54] A. Soualhi, H. Razik, G. Clerc, D.D. Doan, Prognosis of bearing failures using hidden Markov models and the adaptive neuro-fuzzy inference system, IEEE Trans. Ind. Electron. 61 (6) (2013) 2864–2874.
[55] J.B. Ali, B. Chebel-Morello, L. Saidi, S. Malinowski, F. Fnaiech, Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network, Mech. Syst. Signal Process. 56 (2015) 150–172.
[56] S. Saon, T. Hiyama, et al., Predicting remaining useful life of rotating machinery based artificial neural network, Comput. Math. Appl. 60 (4) (2010) 1078–1087.
[57] H. Liao, W. Zhao, H. Guo, Predicting remaining useful life of an individual unit using proportional hazards model and logistic regression model, in: RAMS'06. Annual Reliability and Maintainability Symposium, 2006, IEEE, 2006, pp. 127–132.
[58] A. Widodo, B.-S. Yang, Application of relevance vector machine and survival probability to machine degradation assessment, Expert Syst. Appl. 38 (3) (2011) 2592–2599.
[59] Q. Li, S.Y. Liang, J. Yang, B. Li, Long range dependence prognostics for bearing vibration intensity chaotic time series, Entropy 18 (1) (2016) 23.
[60] Y. Qian, R. Yan, R.X. Gao, A multi-time scale approach to remaining useful life prediction in rolling bearing, Mech. Syst. Signal Process. 83 (2017) 549–567.
[61] H. Qiu, J. Lee, J. Lin, G. Yu, Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics, J. Sound Vib. 289 (4–5) (2006) 1066–1090.
[62] A. Mosallam, K. Medjaher, N. Zerhouni, Time series trending for condition assessment and prognostics, J. Manuf. Technol. Manag. (2014).
[63] B. Zhang, L. Zhang, J. Xu, Degradation feature selection for remaining useful life prediction of rolling element bearings, Qual. Reliab. Eng. Int. 32 (2) (2016) 547–554.
[64] Y. Lei, N. Li, S. Gontarz, J. Lin, S. Radkowski, J. Dybala, A model-based method for remaining useful life prediction of machinery, IEEE Trans. Reliab. 65 (3) (2016) 1314–1326.
[65] Z. Liu, M.J. Zuo, Y. Qin, Remaining useful life prediction of rolling element bearings based on health state assessment, Proc. Inst. Mech. Eng. C 230 (2) (2016) 314–330.
[66] D. Zurita, J.A. Carino, M. Delgado, J.A. Ortega, Distributed neuro-fuzzy feature forecasting approach for condition monitoring, in: Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA), IEEE, 2014, pp. 1–8.
[67] Z. Huang, Z. Xu, X. Ke, W. Wang, Y. Sun, Remaining useful life prediction for an adaptive skew-Wiener process model, Mech. Syst. Signal Process. 87 (2017) 294–306.
[68] Y. Wang, Y. Peng, Y. Zi, X. Jin, K.-L. Tsui, A two-stage data-driven-based prognostic approach for bearing degradation problem, IEEE Trans. Ind. Inform. 12 (3) (2016) 924–932.
[69] L. Xiao, X. Chen, X. Zhang, M. Liu, A novel approach for bearing remaining useful life estimation under neither failure nor suspension histories condition, J. Intell. Manuf. 28 (8) (2017) 1893–1914.

[70] P. Nectoux, R. Gouriveau, K. Medjaher, E. Ramasso, B. Chebel-Morello, N. Zerhouni, C. Varnier, PRONOSTIA: An experimental platform for bearings accelerated degradation tests, in: IEEE International Conference on Prognostics and Health Management, PHM'12, IEEE Catalog Number: CPF12PHM-CDR, 2012, pp. 1–8.
[71] E. Bechhoefer, R. Schlanbusch, Generalized Prognostic Algorithm Implementing Kalman Smoother.
[72] L. Saidi, J.B. Ali, E. Bechhoefer, M. Benbouzid, Wind turbine high-speed shaft bearings health prognosis through a spectral Kurtosis-derived indices and SVR, Appl. Acoust. 120 (2017) 1–8.
[73] L. Saidi, E. Bechhoefer, J.B. Ali, M. Benbouzid, Wind turbine high-speed shaft bearing degradation analysis for run-to-failure testing using spectral kurtosis, in: 2015 16th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA), IEEE, 2015, pp. 267–272.
[74] 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, Mech. Syst. Signal Process. 15 (5) (2001) 945–962.
[75] P. Zimroz, H. Shiri, J. Wodecki, Analysis of the vibro-acoustic data from test rig-comparison of acoustic and vibrational methods, in: IOP Conference Series. Earth and Environmental Science, Vol. 942, IOP Publishing, 2021.
[76] V.C. Leite, J.G.B. da Silva, G.F.C. Veloso, L.E.B. da Silva, G. Lambert-Torres, E.L. Bonaldi, L.E.d.L. de Oliveira, Detection of localized bearing faults in induction machines by spectral kurtosis and envelope analysis of stator current, IEEE Trans. Ind. Electron. 62 (3) (2014) 1855–1865.
[77] H. Shiri, J. Wodecki, Analysis of the sound signal to fault detection of bearings based on Variational Mode Decomposition, in: IOP Conference Series: Earth and Environmental Science, Vol. 942, IOP Publishing, 2021, 012020.
[78] J. Hebda-Sobkowicz, R. Zimroz, M. Pitera, A. Wyłomańska, Informative frequency band selection in the presence of non-Gaussian noise–a novel approach based on the conditional variance statistic with application to bearing fault diagnosis, Mech. Syst. Signal Process. 145 (2020) 106971.
[79] K. Burnecki, A. Wyłomańska, A. Beletskii, V. Gonchar, A. Chechkin, Recognition of stable distribution with Lévy index α close to 2, Phys. Rev. E 85 (5) (2012) 056711.