EPJ AP
Applied Physics
EPJ.org
your physics journal

Eur. Phys. J. Appl. Phys. **93**, 20901 (2021) **DOI:** 10.1051/epjap/2021200259

A neural network approach for improved bearing prognostics of wind turbine generators

Sharaf Eddine Kramti, Jaouher Ben Ali, Lotfi Saidi, Mounir Sayadi, Moez Bouchouicha, and Eric Bechhoefer

edp sciences

The title “The European Physical Journal” is a joint property of EDP Sciences, Società Italiana di Fisica (SIF) and Springer

Eur. Phys. J. Appl. Phys. **93**, 20901 (2021)

https://doi.org/10.1051/epjap/2021200259

THE EUROPEAN PHYSICAL JOURNAL APPLIED PHYSICS

Regular Article

A neural network approach for improved bearing prognostics of wind turbine generators★

Sharaf Eddine Kramti[1], Jaouher Ben Ali[1,2,*], Lotfi Saidi[1,2], Mounir Sayadi[1], Moez Bouchouicha[3], and Eric Bechhoefer[4]

[1] Université de Tunis, ENSIT, Labo. SIME, Av. Taha Hussein, 1008, Tunis, Tunisia.
[2] University of Sousse, ESSTHS-Department of Electronics and Computer Engineering, Sousse 4011, Tunisia
[3] Aix Marseille Univ., Université de Toulon, CNRS, LIS, Toulon, France
[4] Green Power Monitoring Systems Inc. Cornwall, VT 05753, Cornwall, U.S.A.

Received: 8 August 2020 / Received in final form: 10 December 2020 / Accepted: 18 January 2021

Abstract. Condition monitoring of High-Speed Shaft Bearing (HSSB) in Wind Turbine Generators (WTGs) remains a challenging subject for industrial and academic studies. The investigation of mechanical vibration signals presents the most popular method in the literature. Consequently, this work involves a novel data-driven approach for direct HSSB prognosis using the vibration analysis. The proposed method is based on the computation of traditional statistical metrics derived both from the time-domain and frequency-domain via Spectral Kurtosis (SK). Then, the selection of the most suitable features was made using three metrics (monotonicity, trendability, prognosablity) to guarantee a better generalization of the trained Elman Neural Network (ENN). The validation of the proposed method was done using the benchmark of the center for Intelligent Maintenance Systems (IMS) for training and real measured Green Power Monitoring Systems (GPMS) data for testing. We have provided two links for downloading these data sets. The experimental results show that the proposed approach presents a powerful prediction tool. Comparative results with previous work show several advantages for the proposed combination of statistical metrics and ENN, such as the external prediction and real online estimation of the Remaining Useful Life (RUL). Also, some new practical findings are provided in the discussion.

1 Introduction

Many economic factors are based on energy production, especially renewables. Wind Turbine Generator (WTG) is a primary technology for electricity. It is attractive as it has a low carbon footprint and in competitive production cost when compared to natural gas. The most significant negative impact of WTGs technology is the noise, wildlife, visual impact, and the distraction of radar or television reception due to magnetic forces generated by the WTGs and the increased possibility of being struck by lightning. More explanations are mentioned in [1]. Most WTGs have gearboxes and are double-feed induction machines. The controller feeds a rotating three phase field into the stator (1–2 Hz) so that no matter the gearbox speed, the output of the generator is synchronous with the grid. Permanent magnet (PM) machines are asynchronous, the controller picks the speed that is most efficient, and then the voltage is rectified to direct current (DC), then uses an inverter to get grid frequency. Unfortunately, WTGs present several failure modes due to the complicated by high mechanical load and the variable nature of wind. Two extensive surveys of European WTGs over 13 years have identified the origins of WTG failures and the corresponding downtime per failure [2]. Also, statistical studies have demonstrated that WTG gearbox failures account for 10% of all WTG failures [3]. As shown in Figure 1, gearbox failure is the most frequent mechanical failure, the most expensive, and also it presents the highest downtime per failure. According to National Renewable Energy laboratory statistics [4] which are illustrated in Figure 2, the High-Speed Shaft Bearing (HSSB) is most element which have the highest percentage of failure compared to other elements inside the gearbox. Consequently, it leads to sudden downtime, electricity production cessation, and unscheduled maintenance. Hence, Prognostics and Health Management (PHM) of HSSB in WTGs is a primary task that aims to predict the future behavior of the generator's health condition by estimating its Remaining Useful Life (RUL). The ability to predict impending component failure will allow for opportunistic maintenance practices to improve the balance of plant operations and lower maintenance costs.

PHM aims to identify how failure degradation progresses over the next maintenance window and to estimate the remaining time until doing the appropriate maintenance [5,6]. For successful implementation, the

★ Contribution to the Topical Issue "Advanced Materials for Energy Harvesting, Storage, Sensing and Environmental Engineering (ICOME 2019)", edited by Mohammed El Ganaoui, Mohamed El Jouad, Rachid Bennacer, Jean-Michel Nunzi.

* e-mail: benalijaouher@yahoo.fr

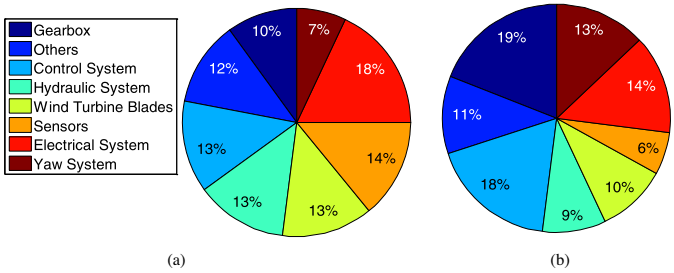


Fig. 1. European WTGs over 13 years: (a) failure/turbine/year, (b) downtime per failure.

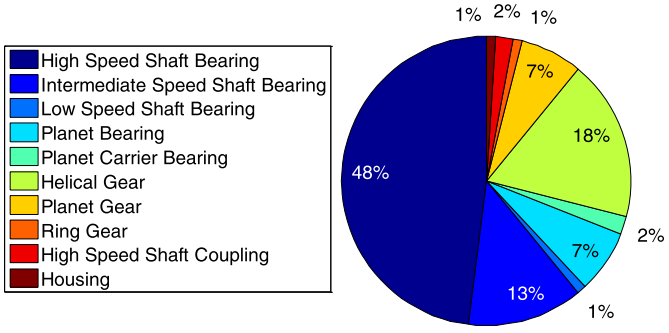


Fig. 2. Failures statistics inside the wind turbine gearbox.

PHM cycle should be composed of three principal parts (norm ISO-13374) ranged as follows:

– observe is the first part that contains two steps: the first one is data acquisition, which collects raw data delivered generally from sensors installed on the desired component. The second step is data processing, which clean, removes noise, and treats the raw data using features extraction and selection algorithms.
– analyze is the second part, which contains three steps. The first one is condition assessment, which uses several metrics or indices to select useful information from the previous step. The second one is the diagnostic that gives the state of component health, fault isolation, and identification according to the previous step. The second step also contains the prognosis, which is planned usually after the diagnosis to estimate the Remaining Useful Life (RUL) and associate the confidence interval of the desired component.
– act is the third part, and it contains two steps. The first one is decision support, where the information is delivered to the decision module, which treats them to create a maintenance intervention. The last step is the human interface machine, which manipulates each part of PHM.

The PHM of HSSB is a vibrant topic in the literature. To estimate accurately the RUL of HSSB based on the investigation of vibration monitoring, a combination of advanced signal processing techniques, artificial intelligence tools, and pattern recognition approaches is needed [7]. In this sense, many researchers have appeared to predict the future health condition of degrading bearing. The next section presents a short literature review.

The remainder of this paper is organized as follows: a summary of bearing RUL based on vibration signals is presented in Section 2. In Section 3, we detail the two data sets used in this work. The different steps of data recording are detailed in this section. Section 4 is devoted to the proposed method and its application for RUL estimation. Section 5 illustrates the experimental results of this paper. A discussion and a comparison with some previous work and methodologies in the literature are given in Section 6. Finally, our conclusions and future work are synthesized in Section 7.

2 Motivation and related works

Mechanical failures in WTGs regularly appeared on bearings drive train due to oil contamination, especially in offshore operations, which caused corrosion in bearings. The varying speed and off-axis gusts result in high loads on bearings, and they do enormous damage / large downtime, resulting in costly maintenance [8]. To maintain balance plant, and power production, it is necessary to estimate the RUL of bearings. As such, several work have been presented in the literature.

The Support Vector Regression (SVR) is widely used for power system applications and is considered a powerful tool [9]. SVR presents a regression model to estimate the relationship between the system input and output from the samples of the training data [10]. However, some parameters of the proposed model need to be carefully defined (the meta-parameters of SVR are the tolerance parameter ε, and the trade-off parameter C. An additional parameter ν can be introduced to deal with the trade-off between the value of ε and model complexity, giving rise to the ν-SVR algorithm) [11]. The accuracy of the estimation depends on the choice of the previous parameters. Therefore, other algorithms are needed to define the model parameters values such as differential evolution algorithm, genetic algorithm, particle swarm optimization, etc. [12]. Hence, the complexity of the problem increases, and the computation burden will be huge. Therefore these solutions tend to remain just theoretical.

The Particle Filter (PF) is also a common approach to bearing RUL estimation. Mainly it is a Bayesian method used for the non-linear process with non-Gaussian additive noise where a physical model is combined with data to present a hybrid approach [13]. To estimate the model parameters, this objective is usually extended to combine parameter and state estimations. Additionally, the PF results depend on a dimension-free error and the number of particles that can be defined after several experiments. Besides, PF lacks validation, and comparative studies between the filters used in the context of prognostics are very scarce and lack of exhaustiveness, which is also true for the prediction procedures [14].

The Gaussian Mixture Models (GMM) has also been used to define a health index that presents better mathematical characteristics. Hence, the bearing RUL estimation task becomes more straightforward, and the developed regression model tends to perform better [15]. The same challenge of model parameter estimation also exists when using the GMM model [16]. Some techniques

can be used, such as negative log-likelihood probability [17].

Unfortunately, any method based on predefined model presents two points of weakness:

– no justification based on physical arguments of the process studied is given when users purpose the model. Even if we try to do that, the number of variables to be considered becomes quite high, which makes the problem very complicated and higher-order.
– the estimation sensitivity of the model parameters is always discarded, and the choice of the estimation algorithm is, unfortunately, dependent on the skills of the user.

Consequently, another alternative is presented in the literature: Artificial Intelligence. This alternative explores computer skills to present more accurate solutions for automatic bearing RUL estimation without human intervention [18]. In [19], a probabilistic RUL prediction was proposed based on a new Recurrent Convolutional Neural Network (RCNN). By considering the temporal dependencies of different bearing degradation states in the uncertainty phase of the RCNN, the RUL prediction uncertainties can be obtained, and a narrow uncertainty interval is achieved. In [20], authors have proposed a novel Recurrent Neural Network (RNN) based on an encoder framework with an attention mechanism. A significant advantage, the proposed neural network was designed to deal with the big data framework. The deep learning algorithm has proved accurate with a long prediction horizon [19,20].

In [18], the authors transformed the prognostic task into a classification task. Seven classes were used before applying six smoothing iterations for the Simplified Fuzzy Adaptive Theory Resonance Map (SFAM) neural network outputs. The learning phase was based on three features fitted using Weibull distribution, and the testing phase demonstrated the ability of the proposed method for bearing failure prognostic.

Before a decade, some researchers have proven the power of neural network for bearing RUL estimation [21–23]. The same architecture was proposed in all these work with the same extraction based tanning phase. In [21], authors have used internal validation (one vibration history was used for training and testing phases). In [18,22,23], authors have overcome this problem using external validation. Some vibration histories were used for training the neural network, and some others were used for testing. With respect, as a weak point, these work have not explained how to do the online normalization of the extracted features. In reality, to do that, we should know the extrema values of features which cannot be done online before reaching the end of the experiment. The output values should also be normalized in the training phase that requires knowing the end date of the experience in advance, and this is the goal and the challenge of prognostic.

Generally, all existing work in the literature investigate the same experimental strategy: the studied signals are recorded using the same test bench, the same bearing model, and under the same experimental conditions. Hence, in this paper, we propose a new methodology based Elman Neural Network (ENN) with a solution for the feature normalization step. Also, we propose to use two different data sets, the first for training and the second for testing. According to our knowledge, there is no previous work that has respected this requirement to generalize the proposed method for a broad range of bearings. In the next section, we present the two data sets used in this work.

3 Brief description of the used methods

3.1 Elman Neural Network (ENN)

Elman backpropagation network named also, Elman Neural Network (ENN) is composed primarily of three layers. The input data has weighted the first layer. The second layer includes typically one or other following hidden layers. Every following layer has weighted from the other layer. All layers have a recurrent weight and bias except for the last one, which is the network output [3]. ENN is believed multi-layer feed-forward neural network with an extremely related layer: it is simultaneous, it fixes various weights, and its training way is based on a back-propagation process. Four procedures are applied in online running:

– amassing inputs data;
– computing hidden unit values;
– computing related values;
– computing output values.

ENN present acceptable accuracy results for time series forecasting thanks to its recursive structure and the connecting layer units. From a theoretical view, the accuracy advantage is ensured via the delayed memory effect of the connecting layer, and consequently, the ENN output values are generally in line with the actual data development trend [24]. Hence, ENN enhances remarkably the sensitivity to historical data compared to the traditional static network, and also it has better dynamic information memory ability [25,26]. The topology structure of ENN is shown in Figure 3.

The operation of the ENN algorithm is primarily based on the hidden layer propagation equation (1), the connecting layer propagation equation (2), and the output layer propagation equation (3).

$$x(k) = F(\alpha_1 x_c(k) + \alpha_2 u(k) + b_1) \tag{1}$$

$$x_c(k) = x(k-1) \tag{2}$$

$$y(k) = G(\alpha_3 x(k) + b_2). \tag{3}$$

We quote that x is the output hidden layer vector, x_c is the output connecting layer vector connecting layer, u is the input layer vector, y is the output layer vector, k is the discrete-time and also known as the sample index, α_1, α_2, and α_3 denote the weight vector respectively between the hidden layer and the connecting layer, between the input layer and the hidden layer and between the hidden layer and the output layer. F and G denote respectively the transfer function of hidden layer and output layer, b_1 and b_2

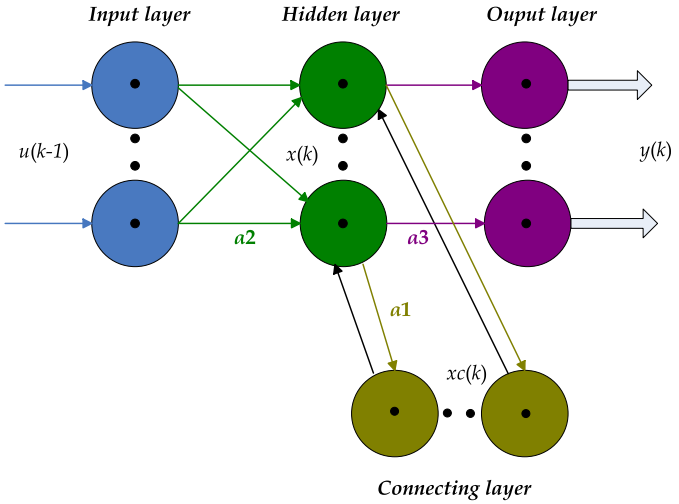


Fig. 3. General topology structure of ENN.

are the bias vectors of hidden layer and output layer respectively. The establishment of the training phase is considered sufficient when the Root Mean Square Error (RMS) will be below a predefined threshold value or after reaching a sufficient number of iterations also set in advance by the user. Mathematically, the RMS is defined by the equation (4) where y_d is the desired output vector and N is the length of each vector (also known as the total number of samples in the training phase).

$$\mathrm{RMSE} = \left(\frac{1}{N}\sum_{k=1}^{N}(y(k) - y_d(k))^2\right)^{1/2}. \tag{4}$$

3.2 Spectral Kurtosis (SK)

Spectral Kurtosis (SK) is a powerful approach with filtering procedure to select the best frequency range and hence to get back demodulated impulsive signatures (that are always hidden in the raw vibrating signal due to noise) [27]. Mathematically, the kurtosis is defined by equation (5) where i is the sample index, x is the vector of samples (the considered discrete signal), x_{mean} and N are respectively the quadratic mean of x and its length (the total number of samples) [7].

$$\mathrm{Kurtosis} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x(i) - x_{\mathrm{mean}})^4}{\left(\frac{1}{N}\sum_{i=1}^{N}(x(i) - x_{\mathrm{mean}})^2\right)^2}. \tag{5}$$

The computation of the SK based on a discrete signal is defined mathematically as the kurtosis of its spectral elements and it's given by equation (6) where X is the cumulant with an exhibitor the presents. The second-order and the fourth-order cumulants are computed using the studied discrete signal x with a band-pass filter around the frequency f and <.> correspond to the time-frequency averaging operator [28].

$$SK(f) = \frac{\langle |X^4(t,f)| \rangle}{\langle X^2(t,f) \rangle^2} - 2. \tag{6}$$

Several advantages lead researchers to use SK especially it is a constant function of frequency in the stationary case and it is identical in the Gaussian stationary case [10]. Nevertheless, bearing vibration signals are considered as non-stationary and affected by quite stronger signals [29]. Consequently, to fully benefit from these advantages, it is recommended to record the mechanical vibrations using high sampling frequency bearings and in a short period to ensure a lower order cyclostationarity.

4 Experimental setup and data recording

4.1 Intelligent Maintenance Systems (IMS) data set

The vibration signals used in this paper were provided by the center for Intelligent Maintenance Systems (IMS), University of Cincinnati, USA, in collaboration with the National Aeronautics and Space Administration (NASA). Four Rexnord ZA-2115 double row bearings are installed on the shaft. Each bearing contains 16 rollers (for each row), pitch diameter of 2.815 in., a roller diameter of 0.331 in., and a tapered contact angle of 15.17° [3,30].

Three tests were made. Each test is an experience of 4 bearings. In this way, 12 bearings are used but only 4 bearings have reached failure with known defects. Each data set describes a run-to-failure experiment. It consists of individual files that are 1-s vibration signal snapshots recorded at specific intervals (every 10min). Each file consists of 20480 points with the sampling rate set at 20 kHz. Records (row) in the data ASCII files are data points. Data collection is provided by NI DAQ Card 6062E [7].

The three tests have been carried out for 35 days until a significant amount of metal debris is found on the magnetic plug of the test bearing [18]. In the first test, two perpendicular sensors (12 o'clock and 3 o'clock positions) were used to record bearing vibration signals. However, only one sensor (12 o'clock) was used. Larger intervals of time stamps (showed in file names) indicate a resumption of the experiment in the next working day. The last file in the first test is uninformative because the record was made in the stoppage phase of the machine [7]. In this paper, we consider only three run-to failure vibration signals that perform a real degradation from the healthy state until the total failure. These three run-to failure experiments are shown in Figure 4. For more details, please visit http://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository/

4.2 Green Power Monitoring Systems (GPMS) data set

The vibration signals are measured from a real-world high-speed shaft bearing installed in real WTG provided by the Green Power Monitoring Systems (GPMS) in the USA.

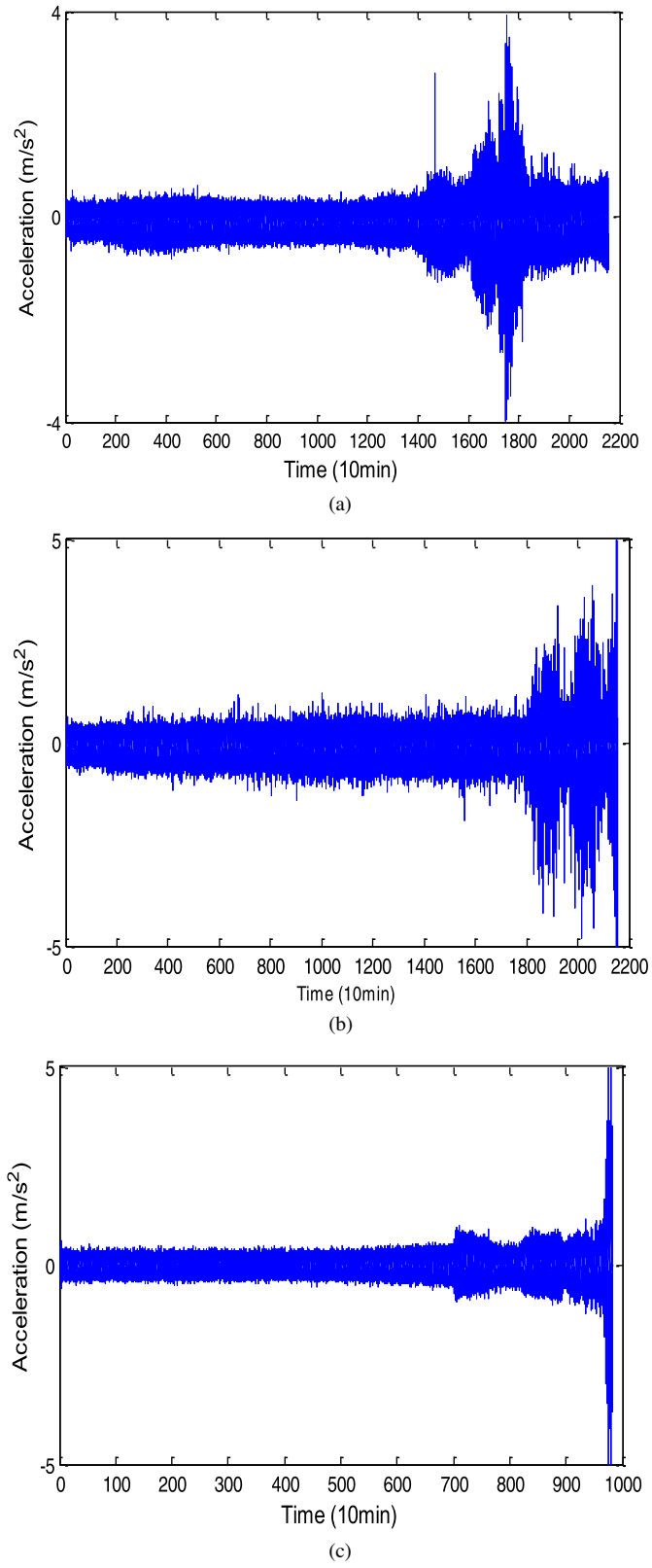


Fig. 4. Run-to-failure vibration signals of (a) bearing 4- testing 1 ending with failure roller. (b) Bearing 3-testing 1 ending with failure inner race. (c) Bearing 1-testing 2 ending with failure outer race.

The data was collected from a real wind turbine generator (S88, Suzelon) with 2 MW electric power generations. The high-speed shaft is trained by thegearbox, the rotation mechanism is assured by 2 bearings installed of the shaft, and the speed was 1800 revolutions per minute during power production. The sample of the radial accelerometer was 100 kHz done for 6 s. More details are provided in [8,31].

The tapered roller bearing (type 32222-J2-SKF) is 200 millimeters (mm) in outer diameter, with a bore of 110 mm and a total length of 56 mm. It has 20 rolling elements at a 16° taper angle and weighs approximately 20 pounds. The operating speed of the bearing is approximately 30 Hz. On the last day, the inspection of the tested bearing has validated an inner race fault.

This data set contains only one run-to-failure experimentwhere the recording time of vibration signals was 6 s per day for 50 days with a high sample rate equal to 100 kHz. The plot of this raw run-to-failure history versus time (days) is shown in Figure 5: it is a simple association of the 50 records, in ascending order of recording days. For more details, please visit https://github.com/mathworks/WindTurbineHighSpeedBearingPrognosis-Data.

5 Experimental results

5.1 Scenario 1: neural network-based bearing prognostic with an internal validation

In this section, we propose to follow the framework given in Figure 6 only for the run-to-failure history of the GPMS data set. Four steps are proposed:

– *feature extraction*: this step consists mainly to compute some traditional statistical features that characterize the degradation of bearing vibration signals with the simplicity of implementation and low central processing unit time [7]. The mathematical expressions of the used nine features in this work are summarized in Table 1. These features are computed, in the first hand in the time-domain, using each vibration record and, in the second hand the frequency-domain, using the SK of each record. Also, the area under the SK is considered as a feature. Thereby, a feature vector with nineteen components is constructed. We quote that x denotes the considered record of the vibration signal, x_{mean} and N are respectively the quadratic mean of x and the length of x, x_{max} and x_{min} are the extrema values of x, k is the sample index and *log* the mathematical operator decimal logarithm. The experimental results of this step are shown in Figures 7 and 8 respectively for the time-domain and frequency-domain.

– *feature selection*: it consists mainly of defining the most significant features for the prognostic task among all ones. This step ensures a better ENN prediction ability and it guarantees a reduced computational time. To meet this constraint, some previous work have investigated the main component evaluation and linear discriminant evaluation [3]. Unfortunately, these approaches generate a new feature set attained after alteration in the new domain that is dissimilar to actual features. To guarantee

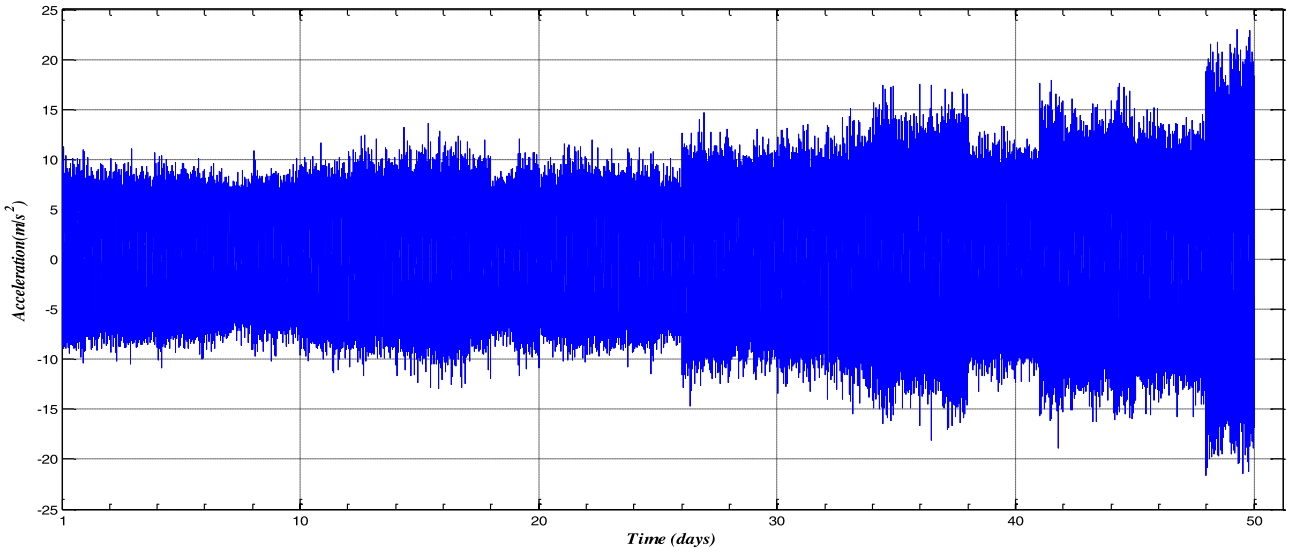


Fig. 5. Raw bearing run-to-failure vibration signals ending with an inner race fault.

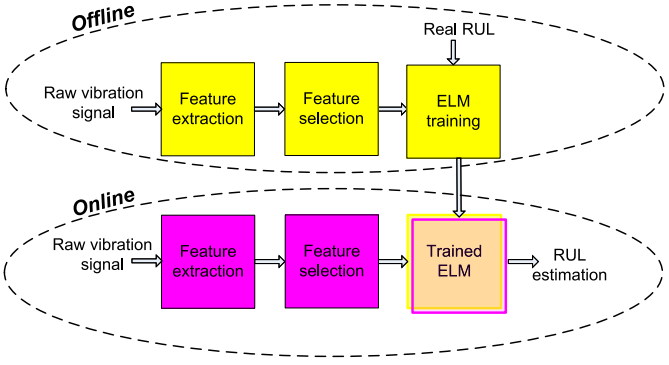


Fig. 6. General synoptic of the first scenario.

a feature selection without modifying them as believed by recent literature on prognosis, there are diverse metrics to distinguish suitable features like trendability, prognosability, and monotonicity [32]. Monotonicity can define the main negative or positive trend of the feature. Theoretically, this is a powerful metric because bearing degradation is an irreversible and growing process. For a discrete signal, the monotonicity of a group of features is affected by the mean between the number of positive and negative step for every assessment point of time. Suppose that n is the number of samples in a particular feature x, the monotonicity is defined by equation (7) [18]. Monotonicity M is measured by absolute difference of *positive* and *negative* derivatives for each feature. The value of M can be from 0 to 1: highly monotonic features will have $M=1$ and non-monotonic features $M=0$. Trendability presents the correlation of the feature in time t. Consequently, it increases with the linearity degree of the feature and it decreases when the feature presents nonlinear characteristics. Particularly, it will be zero when the feature variation is constant. The trendability metric is defined by equation (8) [33]. The prognosability metric is expressed by the formula given in equation (9) is the exponential of the variance of the final failure values divided by the mean range of the path. Features with a high value of prognosability are often ensuring more accurate prognostic [34].

$$\text{Monotonicity} = \left| \frac{\text{number of } dx/dt > 0}{n-1} - \frac{\text{number of } dx/dt < 0}{n-1} \right| \tag{7}$$

$$\text{Trendability} = \left| \frac{n(\sum xt) - (\sum x)(\sum t)}{\left\{ \left[n\sum t^2 - (\sum t)^2 \right] \left[n\sum x^2 - (\sum x)^2 \right] \right\}^{\frac{1}{2}}} \right| \tag{8}$$

$$\text{Prognosability} = \exp\left\{ \frac{-\text{variance}(x_{final})}{\text{mean}|x_{start} - x_{final}|} \right\}. \tag{9}$$

The evaluation of the extracted features based on the proposed three metrics is shown in Figure 9, Tables 2 and 3. The x-axis of Figure 9a and the x-axis of Figure 9b represents respectively the feature index as presented and ordered in the first column of Tables 2 and 3. The y-axis in Figure 9 shows the value of the used metric that should be between 0 and 1. It is really difficult to define the best one for the prognostic. We cannot ensure that a feature with the highest monotonicity value presents also the highest values of trendability and prognosability. Hence, we propose in this work to define a new metric called suitability combining all the previous three metrics together as given in equation (10). The evaluation of the

Table 1. The proposed traditional statistical features.

Feature name	Mathematical expression
RMS	$\left(\frac{1}{N}\sum_{k=1}^{N}x^2(k)\right)^{\frac{1}{2}}$
Mean	$x_{\text{mean}}=\frac{1}{N}\sum_{k=1}^{N}x(k)$
Standard deviation	$std=\left(\frac{1}{N}\sum_{k=1}^{N}(x(k)-x_{\text{mean}})\right)^{\frac{1}{2}}$
Skewness	$\frac{1}{N}\sum_{k=1}^{N}\frac{(x(k)-x_{\text{mean}})^3}{std^3}$
Kurtosis	$\frac{1}{N}\sum_{k=1}^{N}\frac{(x(k)-x_{\text{mean}})^4}{std^4}$
Peak to peak	$x_{\max}-x_{\min}$
Crest factor	$\frac{x_{\max}}{RMS}$
Energy	$\sum_{k=1}^{N}x^2(k)$
Entropy	$-\sum_{k=1}^{N}x(k)\log(x(k))$

suitability feature is provided in Tables 2 and 3.

$$\text{Suitability}=\text{Monotonicity}+\text{Trendability}+\text{Prognosability}. \quad (10)$$

– *ENN training*: the goal of this step is to fix the best ENN architecture for RUL estimation of HSSB. For this, numerous offline tests need to be done until reaching objectives. Unfortunately, the backpropagation neural network structure needs to be compiled many times to confirm a structure that can be not optimal [3]. In the literature there are some empirical recommendations to define the structure of the backpropagation neural network; however, no mathematical justification was given [35]. In this work, we suggest a novel methodology for ENN training. The features that have suitability values higher than the average are selected to load the input layer. Suitability average is shown in Tables 2 and 3 and the selected features that have exceeded the threshold of the average are shown in bold. By this means, 10 features are considered. On the other hand, we added target time also as ENN input. In addition, for each selected feature two values are considered to be given to ENN at the same time: the actual measure and the just prior one. This technique was used a lot of times in previous work and it has proved its efficiency [21,22]. This approach is applied to give to the ENN an impression of curve shape versus time [23]. By this means, 22 ENN inputs are considered (10 features and the target time in the current and the previous times). Besides, we propose to use three hidden layers. The cause of using an ENN with three hidden layers rather than one is to create more interpretative and clear results in proportion to our experiments. The first hidden layer is composed of 15 neurons, the second hidden layer is composed of 7 neurons and the third hidden layer is composed of 3 neurons. The transfer function applied in this ENN is the linear transfer function and the output layer contained one neuron with the saturated linear transfer function using the 0 and 1 thresholds. Consequently, the minimal ENN output value is 0 corresponding to 0% degradation and 1 corresponding to 100% degradation. The target output and the extracted features are normalized using equation (11).

$$\text{feature}=\frac{\text{feature}-\text{feature}_{\min}}{\text{feature}_{\max}-\text{feature}_{\min}}. \quad (11)$$

– *ENN testing*: in this step, the ENN algortithm is used to predict the RUL of HSSB in real condition monitoring with several inputs equal to 20 (without the target time). The dataset which is recorded over 50 days was classified into two subsets: training and testing. The ENN process was training in three modes 20%, 40%, and 60%, of training data. The maximum number of iterations is fixed to 20000. The experimental results of the three modes are shown in Figure 10.

According to Figure 10, the proposed methodology (ENN with gradient descent, momentum and adaptive learning rate static back-propagation of training algorithm) can estimate the HSSB RUL. The predicted RUL is similar to the actual theoretical linear RUL. When we have used 60% of data for training the ENN, the neural network generations have been enhanced and the RUL estimation-becomes more accurate and consequently closer to the linear degradation model.

5.2 Scenario 2: neural network-based bearing prognostic with an external validation

In the previous section, we have repeated the methodology of some previous researches described in [21–23]. Also, we have adopted two major modifications to enhance experimental results. In fact, in the first hand, we have used ENN which is a recurrent neural network instead of using a non-recurrent network and in the second hand, we have selected the best feature instead of using some feature set without any justification. Unfortunately, despite this effort, the internal neural network validation described in the previous section and in [21–23] shows two major points of weakness:

– the validation is internal: in the previous section (and also in [21]) we have used the same bearing for training and testing and this is not possible in a real case. In [22,23] authors have used some bearings for training and

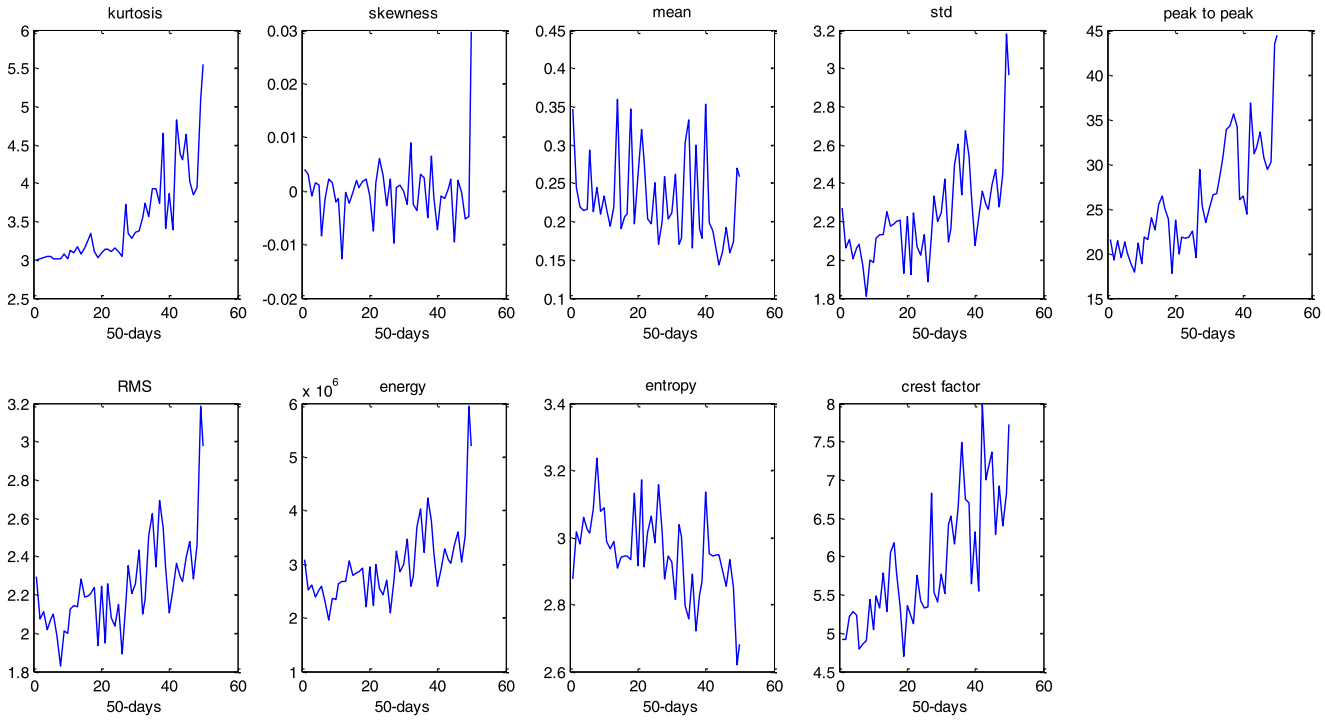


Fig. 7. Time-domain feature extraction results.

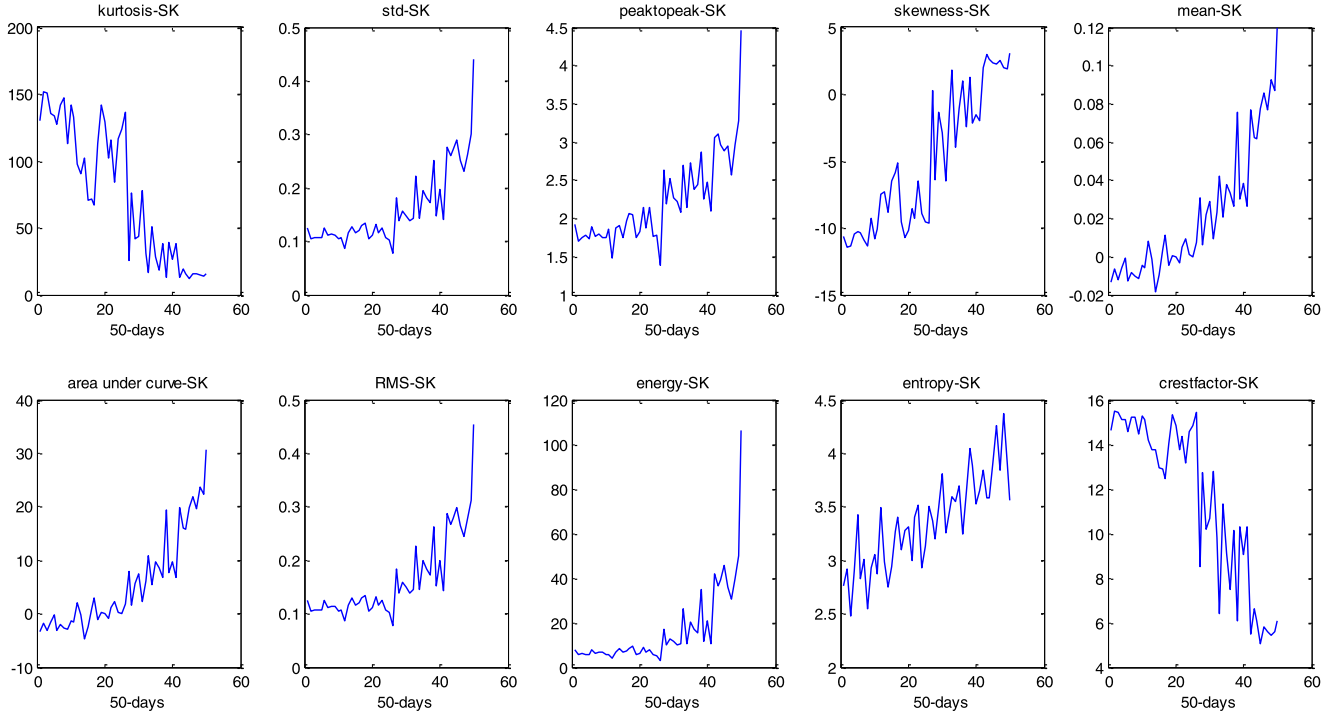


Fig. 8. Frequency-domain feature extraction results based on SK.

other ones for testing. However, this solution cannot be judged as an external RUL estimation as the used vibration signals are recorded from the same bearing model and in same the experimental conditions.

– As shown in Figure 6, some offline steps (with yellow color) and some online steps (with pink color) needs to be done for a neural network methodology. The online step consists mainly of giving the feature vector as input for the validated ENN to estimate the RUL. Since the neural network algorithm is mainly based on mathematical equations and operations, it is necessary to normalize in the same range of numerical values to have the same

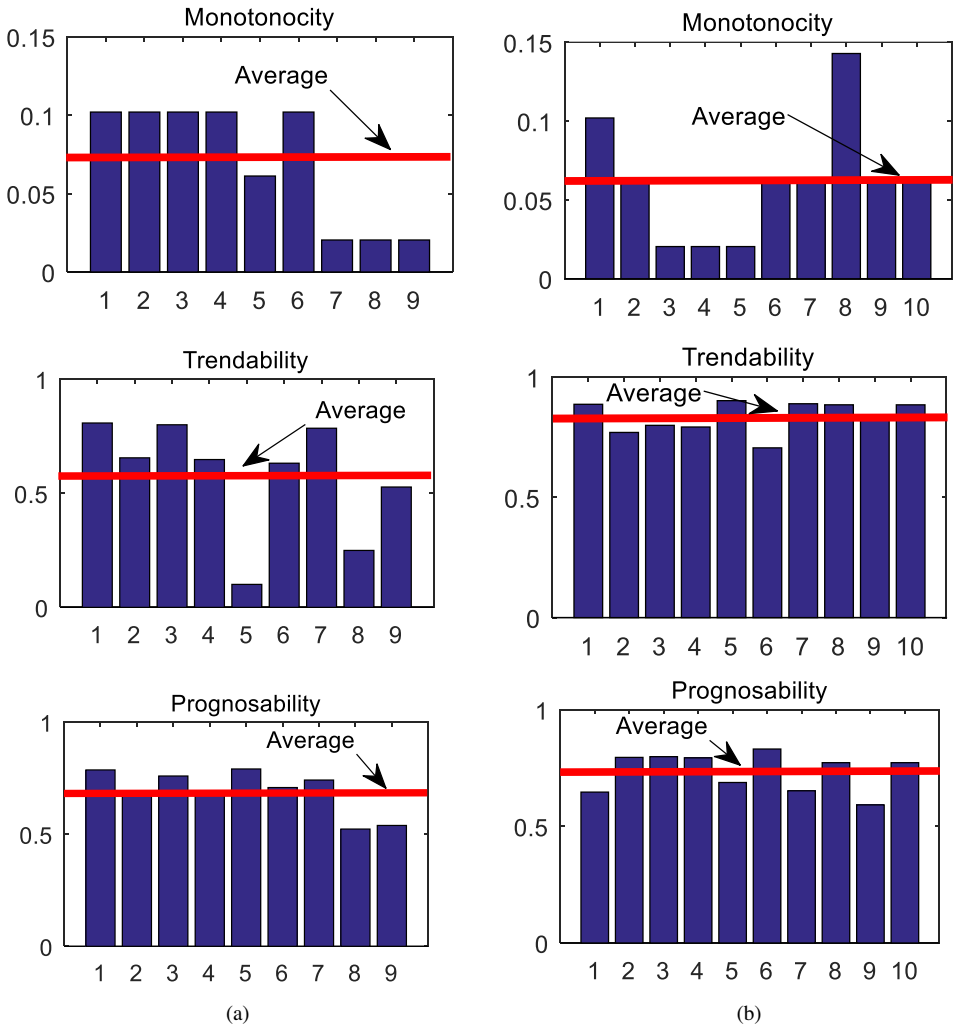


Fig. 9. Evaluation of the extracted features in (a) time-domain, (b) frequency-domain via SK.

Table 2. Proposed traditional statistical features evaluation in time-domain.

Feature index in Figure 8a	Feature	Monotonicity	Trendability	Prognosability	Suitability
1	Kurtosis	0.1020	0.8081	0.7810	**1.6911**
2	Std	0.1020	0.6555	0.6944	**1.4519**
3	Peak to peak	0.1020	0.8004	0.7599	**1.6623**
4	RMS	0.1020	0.6479	0.6892	**1.4391**
5	Skewness	0.0612	0.1007	0.7913	0.9532
6	Energy	0.1020	0.6319	0.7084	**1.4423**
7	Crest factor	0.0204	0.7852	0.7419	**1.5475**
8	Mean	0.0204	0.2496	0.5232	0.7932
9	Entropy	0.0204	0.5275	0.5393	1.0872
All features	Average	0.0703	0.5785	0.6921	**1.3409**

importance in the training phase. In other words, the use of a feature with high numerical values such as the energy feature shown in Figure 7 obliges the ENN to discard other features. That is why a normalization procedure is preliminary. However, to respect the normalization presented by equation (11) we need to know the extreme values of each feature which cannot be done only at the end of the experiment and after the failure of HSSB. Also, we have used the target time as input that cannot be done in real online experiment.

Table 3. Proposed traditional statistical features evaluation in the frequency-domain via SK.

Feature index in Figure 8b	Feature	Monotonicity	Trendability	Prognosability	Suitability
1	Kurtosis-SK	0.1020	0.8873	0.6469	**1.6362**
2	Std-SK	0.0612	0.7702	0.7957	1.6271
3	Peak to peak-SK	0.0204	0.7997	0.7988	1.6189
4	RMS-SK	0.0204	0.7928	0.7938	1.6070
5	Skewness-SK	0.0204	0.9024	0.6881	1.6109
6	Energy-SK	0.0612	0.7061	0.8316	**1.5989**
7	Crest factor-SK	0.0612	0.8896	0.6525	1.6033
8	Mean-SK	0.1429	0.8852	0.7730	**1.8011**
9	Entropy-SK	0.0612	0.8332	0.5920	1.4864
10	Area-under curve	0.0612	0.8851	0.7730	**1.7193**
All features	Average	0.0612	0.8352	0.7345	**1.6309**

In this section, we propose some solutions for the previous mentioned drawbacks. First of all, we have used three bearings of the IMS data set as described in Section 4.1 for ENN training. We have adopted the same ENN architecture validated in the first scenario with a limit of iterations equal to 10^5. As the IMS data set is used offline, it is possible to respect the feature normalization as we know the maximum and the minimum of each feature. Figure 11 shows the shape of the variation over time of some extracted features in the time-domain. Each target output is normalized in the range [0.1] by a simple division by the end time of the experiment. After reaching the ENN training, the online test can be done. First of all, the raw vibration signals of GPMS should be divided by a constant equal 22 to have the same magnitude of the IMS data set. Consequently, 1 s of the first GPMS was used online to estimate the constant value which is 22 and the remaining 5 s will be used to extract features. To ensure the online normalization, and as we have not the extreme value of each feature, the normalization steps of GPMS features will be done online using the maximum and the minimum of the IMS features. All computed features with a value greater than 1 will be considered equal to 1 and all negative ones will be considered equal to 0. The online extracted feature vectors (without the target time) are provided as input for the validated ENN and the RUL prediction results are given in Figure 12. Experimental results are promising and can model real HSSB degradation.

6 Discussion and comparison with some previous works

Concerning previous work and according to our bibliographic investigation, this is the first work that has used two different data sets with different bearing types and use different real experimental conditions: one for training and the other for testing. Therefore, we invite and encourage all researchers to work on this to have a general method for the prognosis of bearings and complete one of the goals of Industry 4.0 challenges. Also, previous work generally divide the same bearing run-to-failure vibration history into two parts: the first one for training and the second one for testing. In [10], authors have used the GPMS data set to validate a RUL method based on SVR and SK. A smaller prediction error was found 60% of training data compared to using 40% of training data. This work is considered as an internal RUL estimation. Moreover, the SVR model parameters estimation was done after reaching 60% of degradation and that cannot be done online. It is impossible to define online the time where the degradation rate is equal to 60%. As well, it is very difficult to build the SVR model and validate it before the recording of the next raw vibration signals. Some applications need to generate RUL estimation each minute (like PRONOSTIA data set) due to the short lifetime of the used bearing and the required precision and perfection (such as in medical and nuclear applications). In other terms, the time between two records is very short compared to the needed time for model parameters estimation ad model validation. In [36] authors have proposed a new Weibull accelerated failure time regression model. Unfortunately, the model parameters estimation needs knowing all features values that cannot be reached online. Therefore, the principal component analysis (PCA) is used for dimensionality reduction. The best principal component value and operating conditions such as speed and load are used in the proposed Weibull model for RUL prediction.

Following the ISO 13381-1, several models can be used for RUL estimation such as heuristic models, statistical models, physics-based models, data-driven models and hybrid models [37]. In [38] a statistical investigation was done on 274 publications has shown that artificial intelligence techniques were used in 81 publications where the neural network has the greatest part with 25 publications. The proposed ENN used in this work is motivated by a feature extraction step to compute some traditional statistical features. These features were dis-

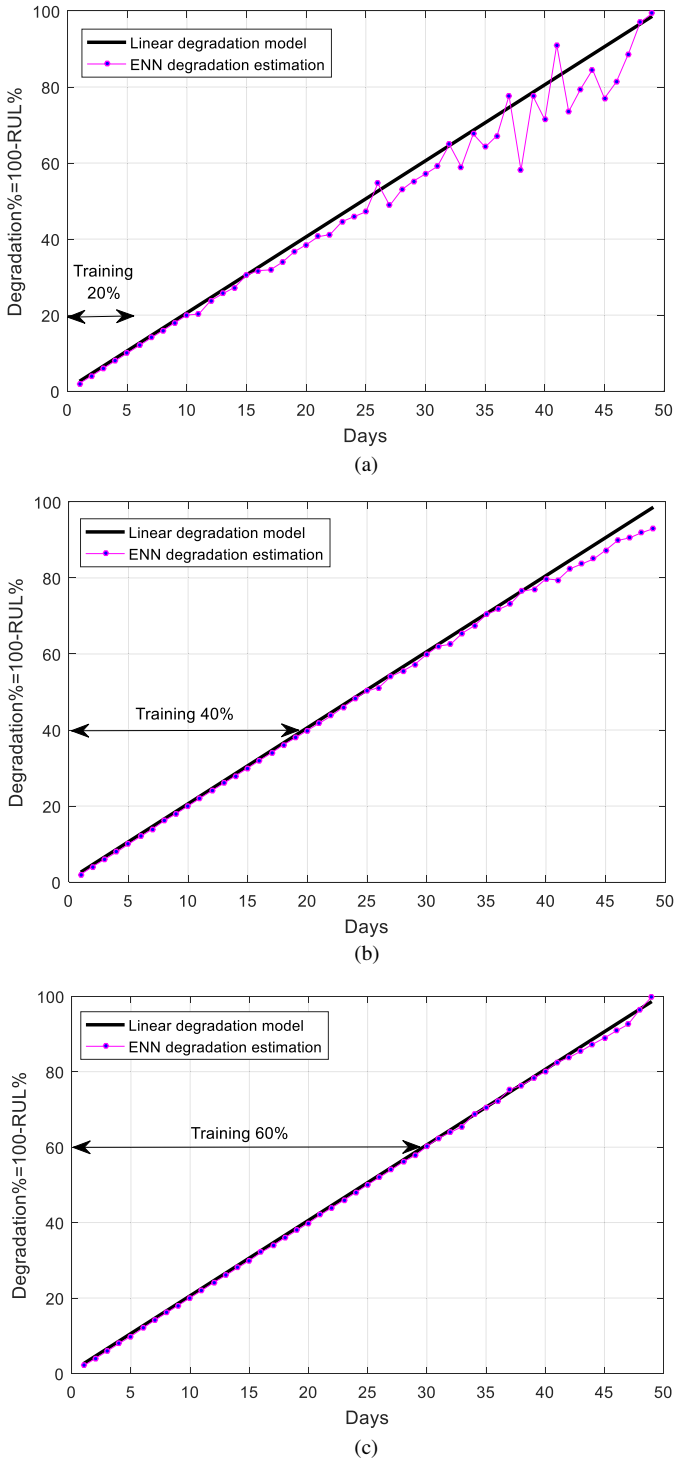


Fig. 10. Internal RUL estimation of HSSB of WTG using (a) 20% of trained data, (b) 40% of trained data, (c) 60% of trained data.

cussed in [39] and authors have concluded that they are redundant for a long time. Consequently, most publications avoid extracting a big number of them as they provide the same information. In this paper, when we have done the feature extraction step, we noticed that redundancy does not present always the same cases. The redundancy exists but it's different from bearing to another. To more explain that, Figure 13 shows that the redundancy is confirmed when comparing RMS and kurtosis features extracted from IMS bearing 1-testing 2. However, this redundancy disappears when using bearings 4 and 3 of testing 1 as shown in Figure 11. Hence, we encourage researchers to use the maximum of features to have a meaningful base that represents carefully the real HSSB degradation.

The feature reduction step is considered important as non-informative features will be then discarded. Thereby, the online computational time will be reduced and an ENN converge can be easier reached. Thus, selecting a suitable feature is a prerequisite for accurate prognostics [38]. All previous work evaluate extracted features in a one-time point using some metrics. However, features evaluation is dynamic. This conclusion is synthesized when having done the online scenario and using the IMS data set, which considered as reach of cases, for ENN training. In fact, by considering the RMS and Kurtosis features extracted from IMS bearing 1-testing 2 in two different points (time = 20 and time = 700), we have found different results. In the first case, the RMS has shown a greater suitability value compared to the kurtosis, but, in the second case, we found the opposite. Henceforth, we invite researchers to define a dynamical feature selection procedure in their next work to ensure more accurate and robust results for HSSB prognostics in WTGs. There is a real need to develop new intelligent methods and to apply different combinations of data collection techniques attempt the prognostic goals [40].

7 Conclusions

In this work, we have introduced the challenge of external HSSB prognostics in WTGs. For this, we have used two different bearing data sets: the IMS data set for training done offline and the GPMS data set for the online testing. Also, we have proposed a new solution for the online feature normalization. Motivated by the good generation of ENN, experimental results were considered as acceptable and encouraging for this challenge. We have proved that the maximum number of extracted features is needed to achieve a big data set that can describe reliably the real degradation of bearings. The number of extracted features would be reduced using some metrics. We have shown in this paper that the step of feature reduction needs to be done with a dynamical behavior. The value of feature criterion metrics depends on the state of HSSB degradation. It generates at each time a different selection from the others. Consequently, we invite the next work to focus on external prognostic and to investigate new methodologies or adopt some existing ones for dynamic feature selection. This ensures more alignment with industry 4.0 requirements.

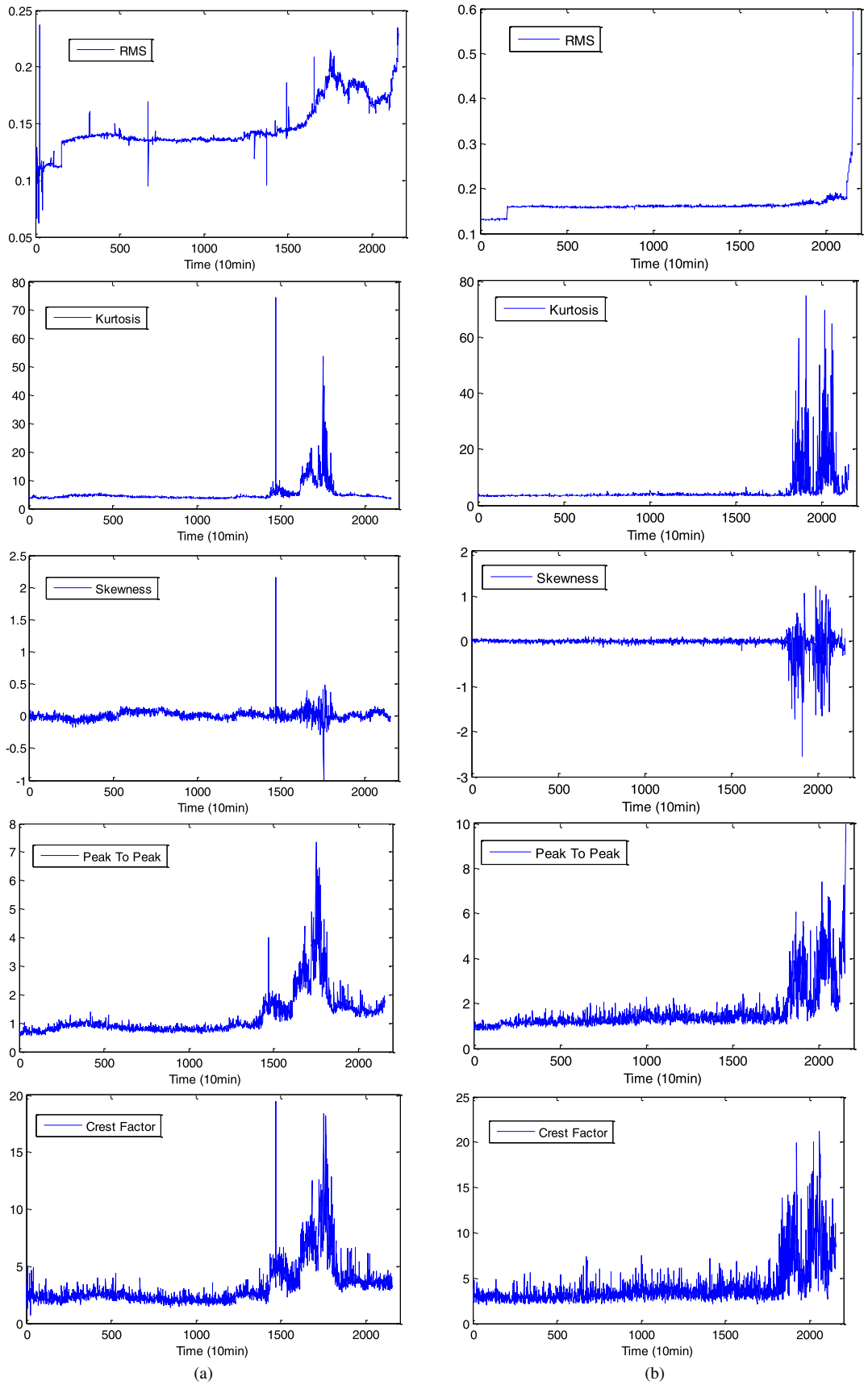


Fig. 11. Examples of extracted feature in time-domain based on IMS data set considering (a) bearing 4- testing 1 ending with failure roller. (b) Bearing 3-testing 1 ending with failure inner race.

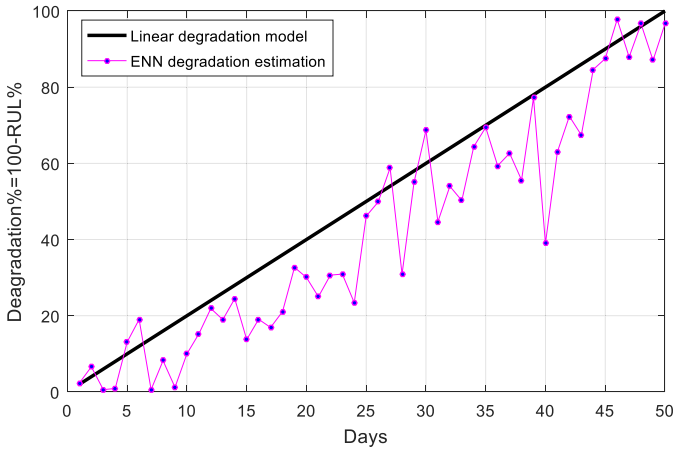
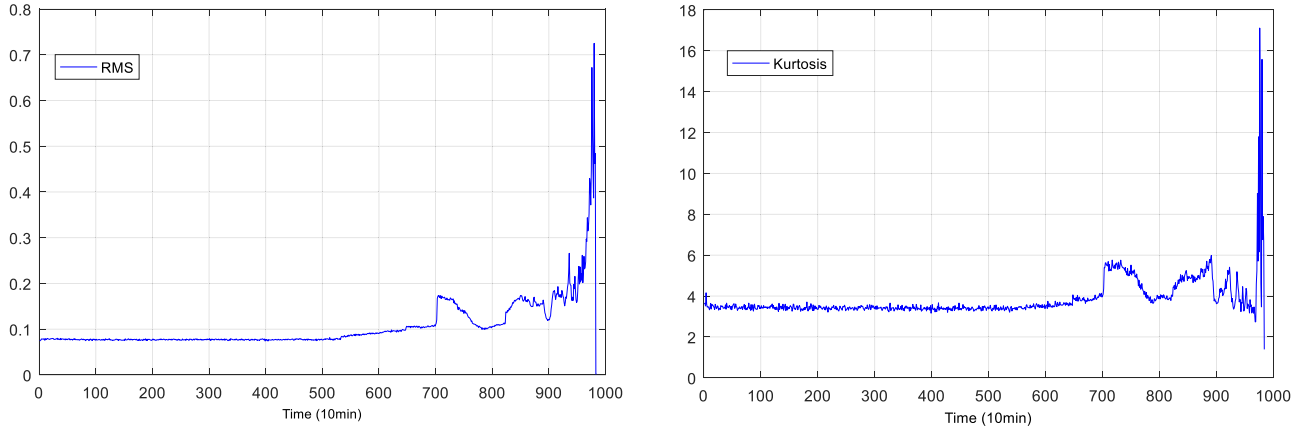


Fig. 12. External RUL estimation of HSSB of WTG.



(a) (b)

Fig. 13. Examples of extracted feature in time-domain of bearing 1-testing 2 ending with failure outer race based on IMS data set.

Author contribution statement

Conceptualization: Sharaf Eddine Kramti and Jaouher Ben Ali.
Methodology: Sharaf Eddine Kramti and Jaouher Ben Ali.
Software: Sharaf Eddine Kramti and Jaouher Ben Ali.
Validation: Sharaf Eddine Kramti, Jaouher Ben Ali and Lotfi Saidi.
Formal analysis: Sharaf Eddine Kramti, Lotfi Saidi, Jaouher Ben Ali; Mounir Sayadi, Moez Bouchouicha, and Eric Bechhoefer.
Investigation: Sharaf Eddine Kramti, Jaouher Ben Ali, Lotfi Saidi and Eric Bechhoefer.
Writing and original draft preparation: Sharaf Eddine Kramti, Lotfi Saidi, Jaouher Ben Ali; Mounir Sayadi, Moez Bouchouicha, and Eric Bechhoefer.
Review and editing: Sharaf Eddine Kramti, Jaouher Ben Ali and Eric Bechhoefer.
All authors have read and agreed to the published version of the manuscript.

References

1. R. Saidur, N.A. Rahim, M.R. Islam, K.H. Solangi, Renew. Sustain. Energy Rev. **15**, 2423 (2011)
2. Y. Zhu, C. Zhu, C. Song, Y. Li, X. Chen, B. Yong, Electr. Power Energy Syst. **113**, 344 (2019)
3. J. Ben Ali, L. Saidi, A. Mouelhi, B. Chebel-Morello, F. Fnaiech, Eng. Appl. Artif. Intell. **42**, 67 (2015)
4. National Renewable Energy Laboratory, "Report on Wind Turbine Subsystem Reliability-A Survey of Various Databases", June, 2013, NREL/PR-5000-59111
5. S.E. Kramti, J. Ben Ali, L. Saidi, M. Sayadi, E. Bechhoefer, Direct wind turbine drivetrain prognosis approach using Elman neural network, in *5th International Conference on Control, Decision and Information Technologies (CoDIT)*, Thessaloniki, Greece, 2018, pp. 859
6. L. Saidi, J. Ben Ali, M. Benbouzid, E. Bechhoefer, Appl. Acous. **138**, 199 (2018)
7. J. Ben Ali, N. Fnaiech, L. Saidi, B. Chebel-Morello, F. Fnaiech, Appl. Acous. **89**, 16 (2015)
8. J. Ben Ali, L. Saidi, S. Harrath, E. Bechhoefer, M. Benbouzid, Appl. Acous. **132**, 167 (2018)

9. A. Soualhi, K. Medjaher, N. Zerhouni, IEEE Trans. Instrum. Meas. **64**, 52 (2014)
10. L. Saidi, J. Ben Ali, E. Bechhoefer, M. Benbouzid, Appl. Acous. **120**, 1 (2017)
11. F. Stulp, O. Sigaud, Neural Netw. **69**, 60 (2015)
12. T. Hamdi, J. Ben Ali, V. Di Costanzo, F. Fnaiech, E. Moreau, J.-M. Ginoux, Biocybern. Biomed. Eng. **38**, 362 (2018)
13. L. Saidi, J. Ben Ali, E. Bechhoefer, M. Benbouzid, Particle filter-based prognostic approach for HighSpeed Shaft Bearing Wind Turbine Progressive Degradations, *43rd Annual Conference of the IEEE Industrial Electronics Society IECON*, 2017, pp. 8099
14. M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, N. Zerhouni, Mech. Syst. Signal Process. **72–73**, 2 (2016)
15. A. Rai, S.H. Upadhyay, Appl. Soft Comput. **71**, 36 (2018)
16. M. Cerrada, R.-V. Sánchez, C. Li, F. Pacheco, D. Cabrera, J. V. de Oliveira, R.E. Vásquez, Mech. Syst. Signal Process. **99**, 169 (2018)
17. J. Yu, Mech. Syst. Signal Process. **25**, 2573 (2011)
18. J. Ben Ali, B. Chebel-Morello, L. Saidi, S. Malinowski, F. Fnaiech, Mech. Syst. Signal Process. **56–57**, 150 (2015)
19. B. Wang, Y. Lei, T. Yan, N. Li, L. Guo, Neurocomputing **28**, 117 (2020)
20. Y. Chen, G. Peng, Z. Zhu, S. Li, Appl. Soft Comput. **86**, 1 (2020)
21. A.K. Mahamad, S. Saon, T. Hiyama, Comput. Math. Appl. **60**, 1078 (2010)
22. T. Zhigang, J. Intell. Manuf. **23**, 227 (2012)
23. T. Zhigang, W. Lorna, S. Nima, Mech. Syst. Signal Process. **24**, 1542 (2010)
24. L. Yang, F. Wang, J. Zhang, W. Ren, Measurement **143**, 27 (2019)
25. K. Kolanowski, A. S wietlicka, R. Kapela, J. Pochmara, A. Rybarczyk, Appl. Math. Comput. **319**, 236 (2018)
26. J. Wang, Future Gener. Comput. Syst. **102**, 670 (2020)
27. J. Antoni, R.B. Randall, Mech. Syst. Signal Process. **20**, 308 (2006)
28. J. Antoni, Mech. Syst. Signal Process. **20**, 282 (2006)
29. R.B. Randall, J. Antoni, Mech. Syst. Signal Process. **25**, 485 (2011)
30. H. Qiu, J. Lee, J. Lin, G. Yu, J. Sound Vib. **289**, 1066 (2006)
31. M. Elforjani, S. Shanbr, E. Bechhoefer, Wind Energy **21**, 1 (2017)
32. C. Jamie, H.J. Wesley, Identifying optimal prognostic parameters from data, *Annual Conference of the Prognostics and Health Management Society*, 2009
33. J. Kamran, G. Rafael, Z. Noureddine, A feature extraction procedure based on trigonometric functions and cumulative descriptors to enhance prognostics modeling, *IEEE Conference on Prognostics and Health Management (PHM)*, 2013
34. S. Harrath, J. Ben Ali, T. Zouaghi, N. Zerhouni, A new adaptive prognostic strategy based on online future evaluation and extended Kalman filtering, in *6th International Conference on Control, Decision and Information Technologies (CoDIT)*, 2019, pp. 2033
35. L. Capocchi, S. Toma, G.A. Capolino, F. Fnaiech, A. Yazidi, Wound-rotor induction generator short-circuit fault classification using a new neural network based on digital data, in *IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics & Drives (SDEMPED), Bologna*, 2011, pp. 638
36. P. Kundu, A.K. Darpe, M.S. Kulkarni, Mech. Syst. Signal Process. **134**, 1 (2019)
37. ISO 13381-1, Condition Monitoring and Diagnostics of Machines-Prognostics-Part 1: General Guidelines, 2015
38. Y. Lei, N. Li, L. Guo, N. Li, T. Yan, J. Lin, Mech. Syst. Signal Process. **104**, 799 (2018)
39. N. Tandon, A. Choudhury, Tribol. Int. **32**, 469 (1999)
40. Z. Liu, L. Zhang, Measurement **149**, 1 (2020)

Cite this article as: Sharaf Eddine Kramti, Jaouher Ben Ali, Lotfi Saidi, Mounir Sayadi, Moez Bouchouicha, Eric Bechhoefer, A neural network approach for improved bearing prognostics of wind turbine generators, Eur. Phys. J. Appl. Phys. **93**, 20901 (2021)