



Estimation of Remaining Useful Life of Rolling Element Bearings Using Wavelet Packet Decomposition and Artificial Neural Network

Abbas Rohani Bastami¹ · Aref Aasi¹ · Hesam Addin Arghand²

Received: 4 May 2017 / Accepted: 12 July 2018 / Published online: 25 July 2018
© Shiraz University 2018

Abstract

Rolling element bearings (REBs) are usually considered among the most critical elements of rotating machines. Therefore, accurate prediction of remaining useful life (RUL) of REBs is a fundamental challenge to improve reliability of the machines. Vibration condition monitoring is the most popular method used for diagnosis of REBs and this is a motivating fact to use recorded vibration data in RUL prediction too. However, it is necessary to extract appropriate features from vibration signal that represent actual damage progress in the REB. In this paper, wavelet packet transform is used to extract signal features and artificial neural network is applied to estimate RUL of the REB. To obtain more accurate results, a method is proposed to find appropriate mother wavelet, optimal level and optimal node for signal decomposition. The desired features were extracted from the decomposed wavelet coefficients. To reduce random fluctuations, which is essential in real-life tests, a preprocessing algorithm was applied on the raw data. A multilayer perceptron neural network was selected and trained by preprocessed input data as well as non-processed input data, and results are compared. A series of accelerated life tests were conducted on a group of radially loaded bearings and vibration signals were acquired in whole life cycle of the tested REBs. Comparison of the experimental results with the output of the trained neural network shows enhanced prediction capability of the proposed method.

Keywords Rolling element bearing · Vibration · Remaining useful life · Wavelet packet transform · Neural network

1 Introduction

Rolling element bearings are widely used in rotating machines and in various industries such as steel, mining, aerospace, paper, railways, textile and renewable energies (Harris 2001). Failure of bearings is a major cause of machinery breakdown, economic losses and even loss of human lives. Undesirable vibrations can be caused by faulty installation, poor maintenance or surface spall that finally leads to development of REB failure (Howard 1994). There is a considerable amount of research in the field of diagnosis of REBs (Singh et al. 2015). One of the

important issues is how to identify bearing fault before it reaches to the final failure state. Bearing failure is reported to be almost 40–50% of the cause of motors failure in industries (Nandi et al. 2005). Based on the critical role of REBs in the machines, it is important to anticipate its RUL in order to take a correct maintenance decision.

Prognostic methods are generally categorized as data-driven methods, model-based methods and hybrid methods (Hu et al. 2012). Data-driven methods are based on processing of vibration signal that is gathered from sensors like accelerometers. Data-driven methods do not require knowledge about the physics of failure. Some of the popular methods used in this category are artificial neural networks (Gebrael and Lawley 2008; Gebrael et al. 2004; Shao and Nezu 2000), regression analysis (Pham and Yang 2010) and neuro-fuzzy approaches (Wang 2007; Liu et al. 2009). Model-based methods by using mathematical models describe the evolution of defect and degradation of REBs like Paris law (Li et al. 1999; Marble and Morton 2006), contact stress analysis (Gebrael and Pan 2008) and

✉ Abbas Rohani Bastami
a_rohani@sbu.ac.ir

¹ Faculty of Mechanical and Energy Engineering, Abbaspour School of Engineering, Shahid Beheshti University, Tehran, Iran

² Mechanical Engineering Department, Sharif University of Technology, Tehran, Iran

Kotzalas and Harris (2001). Model-based methods do not require massive sensory data as is the case in data-driven methods but they are suitable for only special defects where a mathematical model can be applied (Kim et al. 2012). Hybrid methods are combination of data-driven methods and model-based methods (Li et al. 2000; Goebel et al. 2006). Hybrid methods take advantage of both data-driven methods and model-based methods, but similar to model-based methods, hybrid methods are limited to the special defects and cases in which sufficient knowledge is available about physics of the failure.

Heng et al. (2009) published a review paper on the subject of prognosis in rotating machinery which mainly covers prognosis of REBs. After that, two other comprehensive review papers have been published which deal with the prognosis of REBs (Jammu and Kankar 2011; El-Thalji and Jantunen 2015). Si et al. (2011) reviewed statistical data-driven methods for RUL estimation of equipment based on available past data and statistical models. They classified statistical models to two broad types: models that depend on directly observed state information of the equipment and models that do not depend on directly observed state information. They pointed to four challenging problems in RUL prediction that shall be further studied in future works. Qian et al. (2017) used multi-time scale method for prediction of bearing RUL. They divided timescale into fast-time and slow-time scale. Multi-time scale modeling technique called phase space warping (PSW) has been used in damage identification, damage evolution tracking and then RUL of bearing predicted by Paris crack growth model. Qiu et al. (2003) used wavelet filter method to enhance weak signatures of fault and used self-organizing map (SOM) for REB performance degradation assessment. This method is suitable where defect features are impulse-like. Gebraeel et al. (2005) proposed two different exponential degradation model, two-parameter exponential model with multiplicative random error terms and two-parameter exponential model with multiplicative Brownian motion error. They used Bayesian approach for updating estimation of stochastic parameters in exponential models and then used these models for predicting residual life of bearings. Wang and Tsui (2017) proposed a statistical model of bearing degradation. They divided bearing degradation into two distinct stages: In stage 1 bearing stays in normal health condition and also health indicator is in stable level, but in stage 2 bearing defects occur and health indicator exhibits an exponential degradation trend. They used Bayesian inference method from Gebraeel et al. (2005) and extended this method to estimate RUL of bearings more accurately. Guo et al. (2017) estimated RUL of bearings by means of recurrent neural network (RNN) and health indicator. Features for neural network were selected among six related-similarity

(RS) features and eight time–frequency features. The monotonicity and correlation metrics have been used to select the most sensitive fault features. Rai and Upadhyay (2018) predicted RUL of bearing based on self-organizing map (SOM) and support vector regression (SVR). They extracted time-domain and frequency-domain features from the raw bearing vibration signal, and then, by using self-organizing map–minimum quantization error evolution (SOM-MQEE), health index of bearing was achieved. The health index was fed as input to SVR to predict RUL of bearings. Peng et al. (2018) predicted RUL of bearings by using Gaussian mixture model (GMM) and distance evaluation technique (DET). Gaussian mixture model is used to cluster the health states and identify the abnormal data sets and also minimum description length (MDL) method is used to determine the number of clusters. After obtaining the health states, features are chosen based on distance evaluation technique, and then, RUL of the bearings was predicted by means of least square support vector machine (LS-SVM). Huang et al. (2007) used minimum quantization error (MQE) as a degradation indicator from a bearing's incipient defect stage to its final failure stage that is derived from SOM. Then this index is used as an input to a back propagation neural network to predict the residual life of a REB. Hong et al. (2014) mentioned that nature of bearing's vibration signals is non-stationary and has weak faulty signals with strong noise in background; therefore, time–frequency methods are appropriate for feature extraction from the raw signal. They used wavelet packet–empirical mode decomposition (WP-EMD) as a time–frequency method for feature extraction, after that, SOM was used for the condition assessment of the REB performance degradation. Loutas et al. (2013) proposed probabilistic support vector machine (SVM) for prediction of bearing RUL. They used wiener entropy or spectral flatness for condition monitoring of rolling element bearings. Kim et al. (2010) suggested SVM as a classifier to evaluate each health state. They used prior knowledge of the physical degradation of the machinery, failure patterns and information of prior maintenance. Their method needs historical data of various types of defects. Sloukia et al. (2013) used data-driven method for prognosis of REBs. They used SVM as a classification and mixture of Gaussian hidden Markov model (MOG-HMM) for predicting RUL. Liu et al. (2016) classified life cycle of bearings into three health states as normal state, degradation state and failure state. They used data-driven approach for RUL prediction based on multiple health state assessment. They used SVM in classification of health states and also in RUL prediction. Chen et al. (2013) mentioned that SVM is an effective approach for small samples and for univariate time series and to overcome these shortcomings they introduced multivariable support vector machine (MSVM) with

relative features for life prediction. They used features like relative root mean square (RRMS) and relative kurtosis factor as input; they concluded that MSVM has a better performance than univariate SVM in RUL prediction when insufficient condition monitoring data exist. They concluded that complicated signal processing methods are required to assess the REB performance degradation and determine the initial defect as well as the final failure more accurately. Reuben and Mba (2014) used actual service data gathered from AH64D helicopters that was measured by accelerometer and used spectral analysis of the data in prognosis. They claimed that in both low- and high-frequency bandwidths, simple spectral analysis was effective for tracking progressive stages of bearing damage. They used regression models like exponential regression model and 5-parameter logistic model for RUL prediction. Ali et al. (2015) defined a new feature as root mean square entropy estimator (RMSEE) and used simplified fuzzy adaptive resonance theory map (SFAM), neural network and Weibull distribution (WD) to calculate RUL of REBs. Tian et al. (2010) used suspension data whereby machines are taken out of service before they fail, in addition to condition monitoring data from failure histories, and used them as inputs to a feedforward neural network (FFNN) for predicting RUL. They concluded that degradation process is important for correct prediction of RUL. Zhang et al. (2013) decomposed vibration signal into several signals containing one approximation and some details by wavelet transform (WT) and then transformed them to frequency domain using fast Fourier transform (FFT), and then they were used as inputs to train a back propagation neural network.

In this paper, a novel prognostic method is proposed that uses current and past vibration signal of the machine. In the proposed method, vibration signal is decomposed by wavelet packet transform (WPT). To obtain better result, best mother wavelet has been selected among 53 examined wavelet bases. Then optimal level and node have been chosen based on kurtosis value. Then features are extracted from decomposed signal and are fed to a designed artificial neural network. In this paper, multilayer perceptron (MLP) network with preprocessing data is employed for RUL prediction. The proposed method showed good accuracy in the prediction of RUL using data obtained in REB life test and proved its capability for practical applications.

This paper is organized as follows. In Sect. 2, the proposed method is explained. Experimental setup is described in Sect. 3. Section 4 explains wavelet transform and common criteria for selection of appropriate mother wavelet, optimal level and node. In Sect. 5, artificial neural networks are explained, and results of MLP are compared with experimental results. Conclusions are given in Sect. 6.

2 Method

Data-driven prognostic methods do not need to model evolution of defect and physics of defect growth. This method is based only on measured vibration signals. In this paper, we used wavelet packet transform and ANN to predict RUL of a deep groove ball bearing under controlled load. Schematic diagram of the proposed method is shown in Fig. 1.

Feature extraction is based on presentation of vibration signal in different domains. In time domain, main features are based on statistical description of the vibration signal like peak, root mean square (RMS), crest factor, kurtosis. (Chen et al. 2013). Common time-domain features are defined in Table 1. Other features are calculated based on transformation of the signal into other domains like frequency domain by means of Fourier transform. However, simple FFT is not efficient in analysis of non-stationary signals, and time–frequency-domain features are supposed to provide more information. The main methods to transform signal into the time–frequency domain are short time Fourier transform (STFT) and wavelet transform (WT). STFT has problems in selecting window length and time–frequency resolution. WT has a compact dyadic presentation of the signal and is a good choice for data compressing, while WPT has more detailed information. Therefore, WPT has been selected for extracting appropriate features of REB vibration signal.

Features in Table 1 can be described as follows: Mean is mathematical expectation of the random variable. RMS represents the average power of a signal. Absolute average value is used to get the absolute values of all the numbers, and then, average function is applied to calculate the result. Fourth moment is a specific quantitative measure, used in

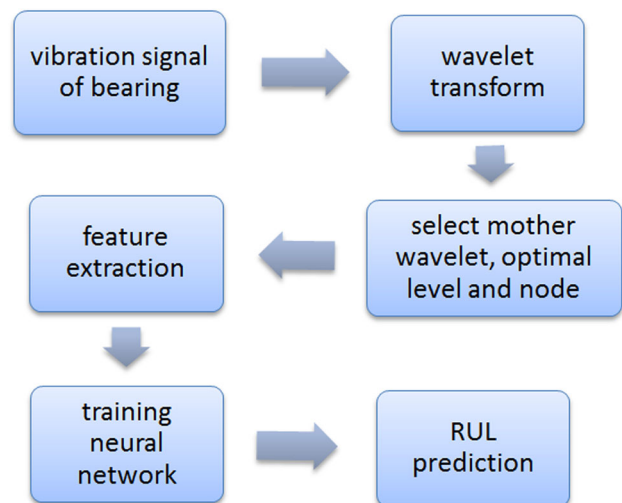


Fig. 1 Process of the proposed REB prognosis method

Table 1 Common statistical vibration features in time domain

No.	Feature	Formula
1	Mean	
2	Root mean square (RMS)	$RMS = \sqrt{\frac{1}{M} \sum_{k=1}^M x^2(k)}$
3	Absolute average value	$\bar{x} = \frac{1}{M} \sum_{k=1}^M x(k) $
4	Fourth moment	$\beta = \frac{1}{M} \sum_{k=1}^M (x(k) - \mu_x)^4$
5	Variance	$\sigma_x^2 = \frac{1}{M-1} \sum_{k=1}^M (x(k) - \mu_x)^2$
6	Peak value	$x_p = \max(x(k))$
7	Crest factor	$C = \frac{x_p}{x_{rms}}$
8	Kurtosis factor	$K = \frac{\beta}{\sigma_x^4}$

Where for $k = 1, \dots, M$, $x(k)$ is a discrete signal and k is the number of sequence

both mechanics and statistics of the shape of a set of points. Variance is the expectation of the squared deviation of a random variable from its mean, informally, it measures how far a set of random numbers are spread out from their average value. Peak is the maximum value, either positive or negative, that a waveform attains. Crest factor is the ratio of peak values to the RMS. In other words, crest factor indicates how extreme the peaks are in a waveform. Kurtosis is a measure of the tailedness of the probability distribution of a random variable in comparison with the normal distribution.

3 Experimental Setup

A test rig for accelerated life test of REBs at Sharif University of Technology is used to verify proposed prognosis method. The schematic of test rig is shown in Fig. 2. The vibrations data of each bearing consist of two channels from two B&K AS-65 accelerometers that are placed radially on the housing of the test bearing. The test bearing type is NTN 6804-ZZ, a double-shielded single-row deep groove REB. The experimental setup is shown in Fig. 3.

As shown in Fig. 3, two accelerometers were mounted in vertical and horizontal directions. The load was applied to the bearing radially in vertical direction by means of four screws and measured by a load indicator. Bearings are shielded so there is no need for external lubrication. Vibration signal was collected at two sampling frequencies of 51.2 kHz by an eight channel, 16 bits real-time data recorder and 25.6 kHz by a portable two channel vibration analyzer. In this experiment, six bearings have been tested as shown in Table 2. All tests have been done at constant speed (1500 rpm) and approximately constant radial load (4400 N) with a small variation. Bearing life depends on load but has stochastic behavior. Load changes in the tests were small; therefore, variation in bearings life is due to randomness in bearing life. According to the conducted tests and observations, acceleration threshold for end of the test was selected as 20 g rms. Inspection of internal parts of the bearing showed that bearing damage was fully developed at this acceleration level. Since run of the test beyond this level would have safety risks, this acceleration level was chosen as end of test. This acceleration threshold has been reported in other tests with the same size of bearing (Wang et al. 2015; Zhao et al. 2016). In addition, temperature of the bearings was measured in every 15 min by a thermometer during the test.

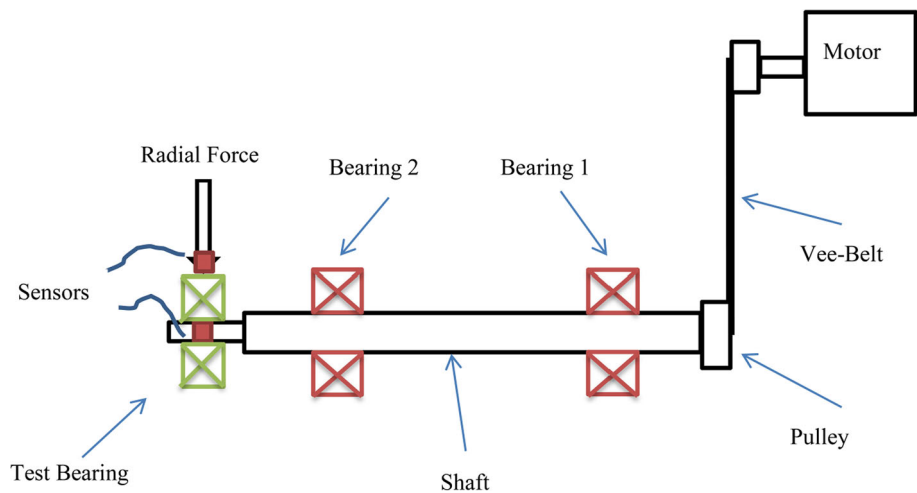
Fig. 2 Schematic of test rig

Fig. 3 Experimental setup for REB run-to-failure test

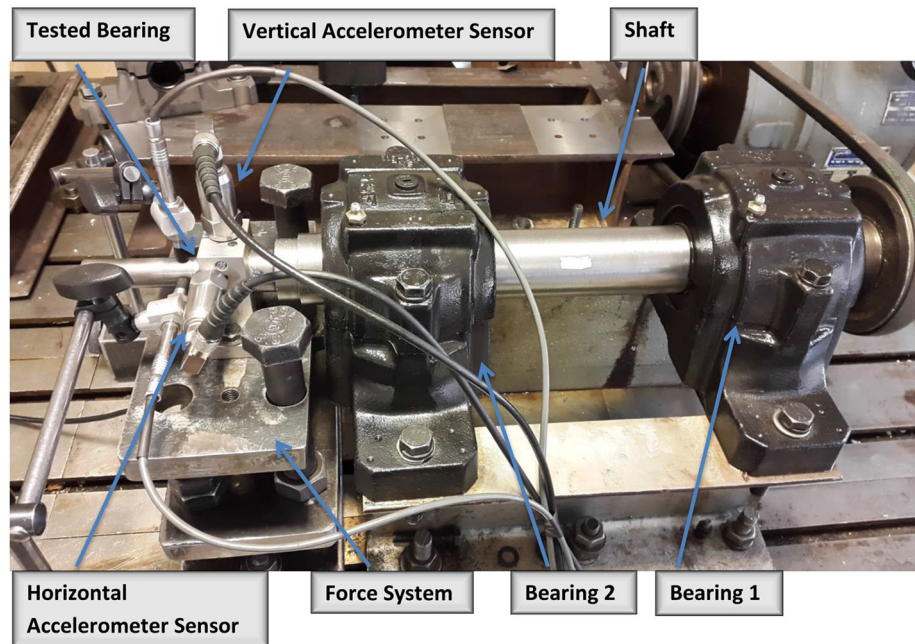


Table 2 Test specifications

Tested bearing	Load (N)	Speed (rpm)	Life (min)
Data test 1	4312	1500	471
Data test 2	4410	1500	624
Data test 3	4410	1500	741
Data test 4	4459	1500	986
Data test 5	4459	1500	1200
Data test 6	4459	1500	1862

The run-to-failure measured vibration signal of bearing test 6 is shown in Fig. 4. As shown in Fig. 4, amplitude of

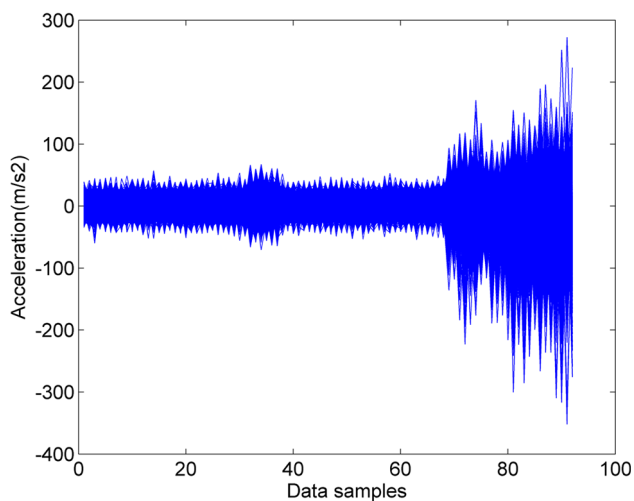


Fig. 4 Bearing vibration signal in test no. 6

the acceleration increases until it reaches the threshold of 20 g. It shall be noted that at this acceleration level bearing defect has been grown entirely and in real-world applications the machine must be stopped earlier to avoid damage to other parts. Nevertheless, the test was continued to reach this level to observe defect growth in the bearing until near the end of the bearing life. Last portion of the RMS trend of vibration in bearing test 5 is plotted in Fig. 5. Damaged bearing in test no. 5 is shown in Fig. 6. Spalls can be seen on the inner race and balls. In real applications, the threshold for stopping the machine shall be below 20 g rms to avoid safety risks.

4 Wavelet Transform (WT)

Classical diagnosis tools like fast Fourier transform (FFT) that are used for the analysis of stationary vibration signals are not suitable for analysis of bearing defects (Akbari et al. 2014). Vibration of defective bearings can be considered as cyclostationary (Antoni 2009), and time–frequency or bi-frequency transforms (Yiakopoulos and Antoniadis 2005) are more suitable for it. Wavelet theory developed as a signal processing tool which is suitable for analysis of both stationary and non-stationary signals. This transform has also been used successfully in diagnosis of bearings and feature extraction (Kankar et al. 2011). The commonly used wavelet transforms are continuous wavelet transform (CWT), discrete wavelet transform (DWT) and wavelet packet transform (WPT) (Kulkarni and Sahasrabudhe 2013).

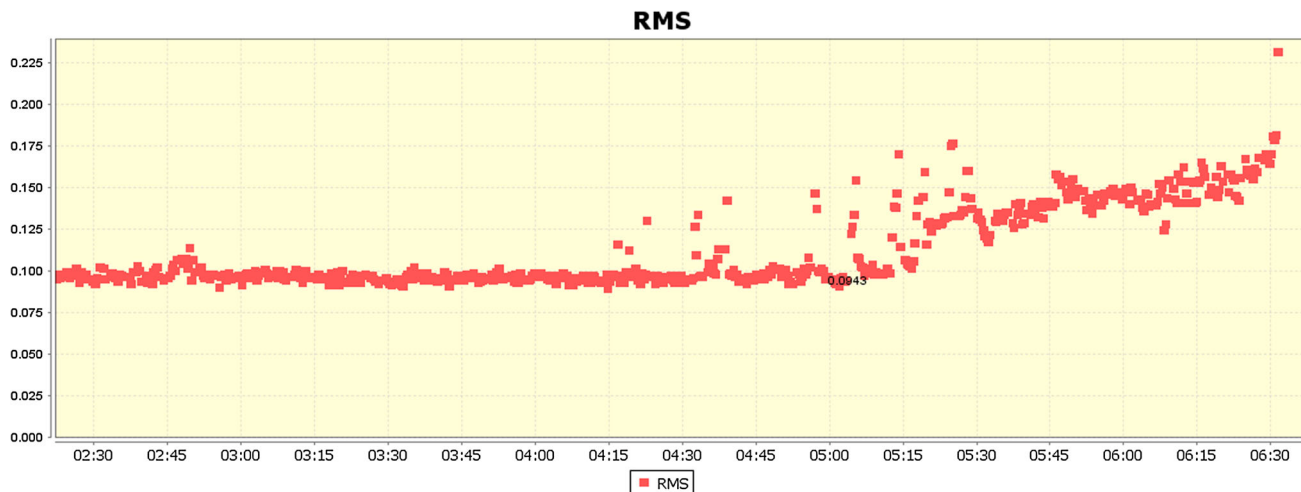


Fig. 5 Trend of acceleration RMS (km/s²) versus hours of operation in bearing test no. 5



Fig. 6 Inner race and rolling elements of the damaged bearing in test no. 5

CWT generates redundant data, and its calculation consumes more time than other wavelet transforms. In comparison, DWT is more compact and faster. This is the reason that DWT is more common in condition monitoring purpose (Akbari et al. 2014).

Continuous wavelet transform of signal $x(t)$ is defined as:

$$T(a, b) = \int_{-\infty}^{+\infty} x(t)\psi_{(a,b)}^*(t)dt \tag{1}$$

The sign * is complex conjugate and $\psi_{(a,b)}$ is:

$$\psi_{(a,b)}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \tag{2}$$

$\psi_{(a,b)}(t)$ is a daughter wavelet that is obtained from mother wavelet, a and b are real parameters, representing scaling and translation, respectively.

Discrete wavelet transform is obtained by discretization of CWT. The most popular one is dyadic discretization given by:

$$\psi_{(j,k)}(t) = \frac{1}{\sqrt{2^j}}\psi\left(\frac{t-2^jk}{2^j}\right) \tag{3}$$

where a and b are replaced by 2^j and 2^jk , respectively.

Wavelet packet transform is a generalized form of DWT where in addition to approximation coefficients, detail coefficients are also decomposed and all nodes at the same level have equal frequency bandwidth. WPT is calculated by `wpdec` command in MATLAB.

4.1 Selection of Best Mother Wavelet and Optimal Node

There are various methods for selecting mother wavelet and decomposition level. Yan (2007) proposed maximum energy to Shannon entropy ratio as a criterion for mother wavelet selection. Rafiee et al. (2009) used Daubechies (DB) mother wavelet and have chosen 4th level as an optimal level. Kankar et al. (2011) extracted statistical features by using complex Morlet wavelet based on minimum Shannon entropy criterion (MSEC). Bafroui and Ohadi (2014) decomposed vibration signals of a gearbox by using CWT and Morlet wavelet and determined the optimal range of wavelet scales based on the maximum energy to Shannon entropy ratio criterion. Kumar et al. (2014) used the fact that mother wavelet selection depends on the similarity between the shape of original signal and the mother wavelet and also focused on the mother wavelet selection for analyzing bearing vibration signals based on the two criteria including MSEC and maximum energy to Shannon entropy ratio criteria. Akbari et al. (2014) chose appropriate mother wavelet and level based on maximum energy to Shannon entropy ratio criteria.

Combination of two criteria of maximum energy and minimum Shannon entropy for mother wavelet selection leads to the maximum energy to Shannon entropy ratio criteria (Bafroui and Ohadi 2014):

$$\xi(n) = \frac{E(n)}{S_{\text{entropy}}(n)} \quad (4)$$

That $E(n)$ is energy at level n and $S_{\text{entropy}}(n)$ is entropy at level n .

Energy at level n is given by:

$$E(n) = \sum_{i=1}^m |C_{n,i}|^2 \quad (5)$$

where $C_{n,i}$ is the i th wavelet coefficient at level n , m is the number of wavelet coefficients.

Shannon entropy at level n is given by:

$$S_{\text{entropy}}(n) = - \sum_{i=1}^m P_i * \log P_i \quad (6)$$

where P_i is the distribution of the energy probability for the wavelet coefficients given by:

$$P_i = \frac{|C_{n,i}|^2}{E(n)} \quad (7)$$

where $\sum_{i=1}^m P_i = 1$. The value of $P_i * \log P_i$ is assumed zero when $P_i = 0$.

In this paper, maximum energy to Shannon entropy ratio criteria have been used for mother wavelet selection. Comparing 53 different mother wavelets as illustrated in

Fig. 7 shows that bior 3.1 is the most suitable mother wavelet for REB feature extraction.

Due to the fact that WPT is a generalization of DWT and provides richer information of the considered signal (Nikolaou and Antoniadis 2002; Peng and Chu 2004; Al-Badour et al. 2011), WPT has been used in this paper and both optimal level and optimal node are selected.

Kurtosis of acceleration signal is used widely in diagnosis of rolling bearings. Healthy bearing has a kurtosis of 3, while higher kurtosis values correspond to bearing defect. To select the optimal level and node, the value of kurtosis of decomposed signal in all nodes of each level is calculated. The node that has a larger value is selected as the optimal level and the optimal node.

By applying WPT on the bearing vibration signal and calculating the value of kurtosis at each node, as shown in Fig. 8, Node (4, 7), i.e., level 4 and node 7, are chosen as an optimal level and an optimal node. Two features, RMS and Kurtosis, are calculated from decomposed signal at this node and fed as input of the neural network.

5 Artificial Neural Network (ANN)

ANN structure consists of one input layer, one or more hidden layer, one output layer, weights and biases. Process of determining the weights and the biases is known as training of the ANN (Rao et al. 2012). Schematic diagram of the multilayer ANN is shown in Fig. 9. The extracted features from the vibration signal act as network inputs, and target or output is the RUL of the bearing stated in percentage of total life. Selection of the features for ANN is based on the comparison of performance of eight features that are listed in Table 1. Root mean square of error (RMSE) is used to measure performance of each feature. Performance of the ANN for all combinations of inputs in groups of one, two and three features was calculated. The best features as single input were used in combination to form two and three inputs. Best results were obtained in groups of two features as shown in Table 3. It can be seen that combination of RMS and absolute average value has less error than others and consequently this combination was chosen as input to ANN.

A MLP neural network with seven neurons (chosen by comparing the results) in the hidden layer and one neuron in the output layer is used. Hyperbolic tangent sigmoid transfer function is selected for hidden layer and linear transfer function for output layer. Levenberg–Marquardt algorithm is used for ANN training, because it gives better performance among other network training algorithms. Cost function is selected as mean square of errors (MSE). The NN has been trained by RMS and absolute average value as inputs and the percentage of RUL as target. There

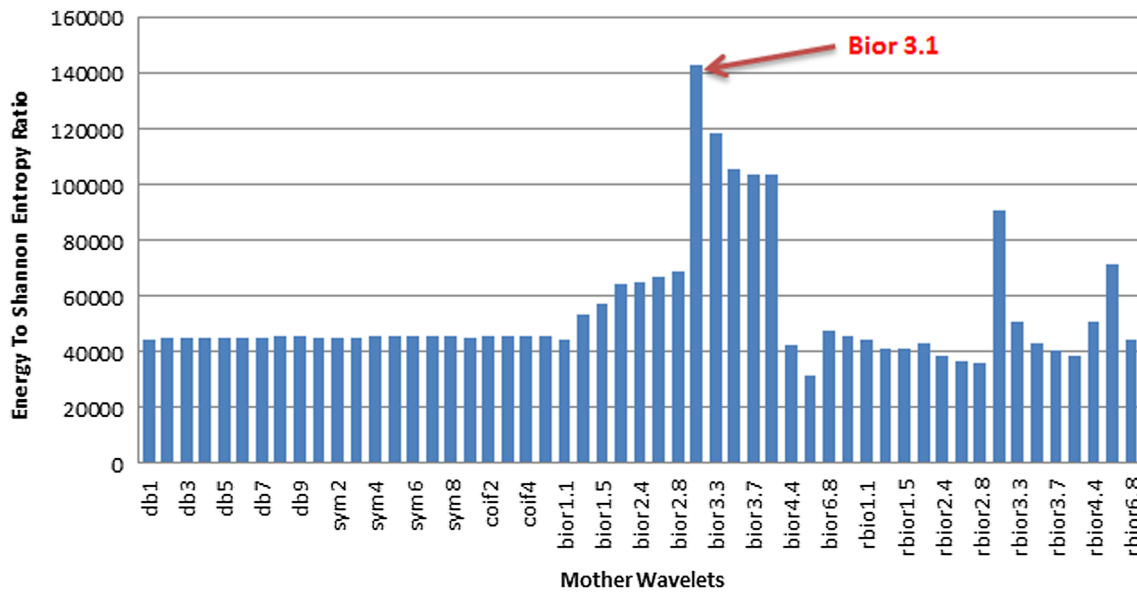


Fig. 7 Comparison of energy to Shannon entropy ratio of different mother wavelets

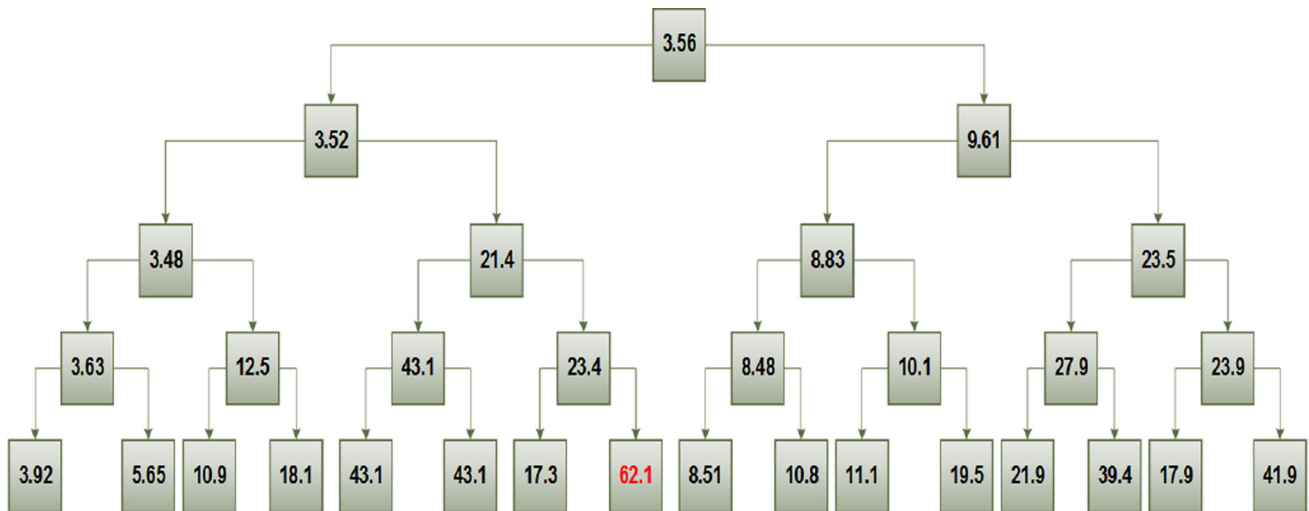


Fig. 8 Optimal level and optimal node selection based on kurtosis value

were 487 data points in which 347 data points are selected randomly for training, and the remaining 70 random data points are used for validation and test. Result of the MLP for data test 5 is shown in Fig. 10.

The prediction result obtained by MLP is not satisfactory. To improve the accuracy of MLP prediction, some preprocessing has been applied on the features obtained from decomposed signal by WPT. First, input data were smoothed by moving average method. The smoothed RMS is depicted in Fig. 11. It can be seen that RMS has nearly monotonic trend so it can be used in prognosis. Features that do not have monotonic behavior such as kurtosis are not good candidates for prognosis. Then, MLP has been

trained by smoothed data, and the results as depicted in Fig. 12 show better behavior.

To further improve results, after smoothing of inputs, a Weibull basis hazard function $h(t)$ and exponential function were fitted to them. Weibull hazard function $h(t)$ is defined as (Jardine et al. 1998):

$$h(t) = \frac{\text{probability density function (PDF)}}{1 - \text{cumulative distribution function (CDF)}} \quad (8)$$

$$h(t) = \frac{\beta}{\eta^\beta} t^{\beta-1} \quad (9)$$

where β and η are shape and scale parameters of Weibull distribution function, respectively. The Weibull shape parameter β is also known as the Weibull slope because the

Fig. 9 Schematic diagram of the multilayer ANN

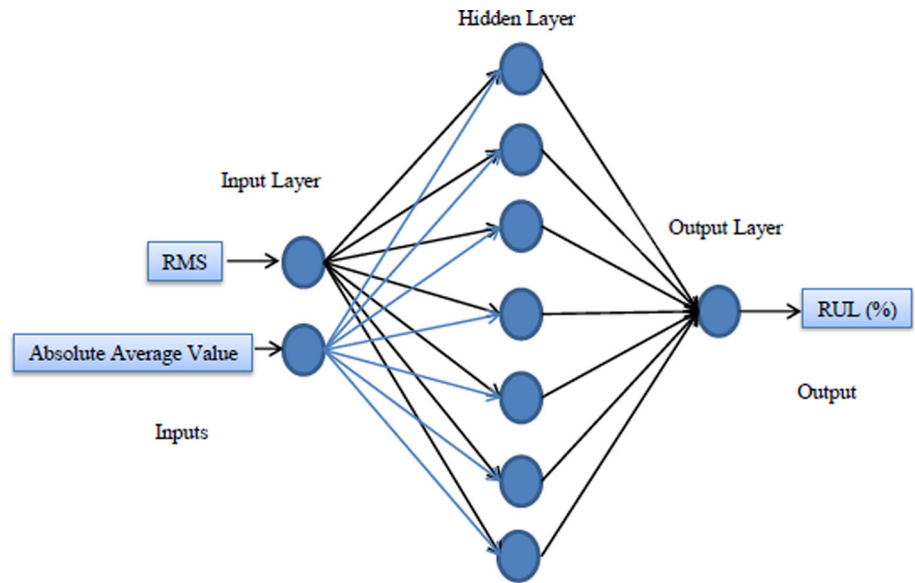


Table 3 Comparison of RMSE of various combinations of vibration features as input to ANN

Features	Root mean square of error, RMSE (%)
Absolute average value + four moments	3.49
Absolute average value + KURTOSIS	2.26
Absolute average value + RMS	0.98
Four moments + KURTOSIS	3.52
RMS + four moments	3.42
RMS + KURTOSIS	2.13

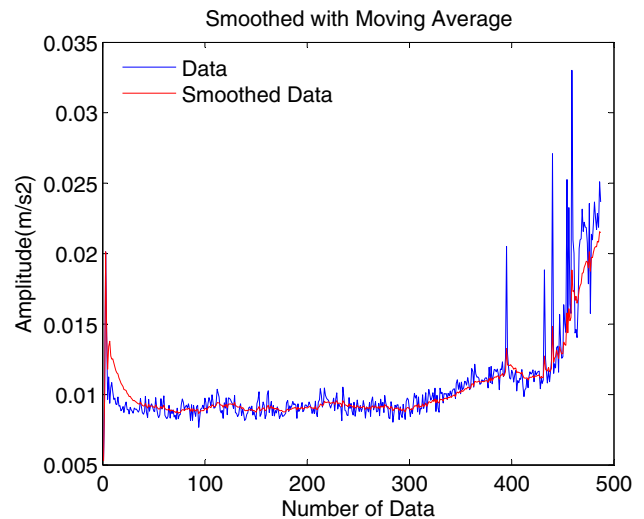


Fig. 11 RMS was smoothed with moving average method

value of β determines the initial slope of the curve. Change of scale parameter η is equivalent to the change in the abscissa scale.

Exponential function $f(t)$ is defined as:

$$f(t) = ae^{bt} + ce^{dt} \tag{10}$$

Fitting of RMS by Weibull basis hazard function is depicted in Fig. 13, the R -Square of this fitting is 0.7, and its root mean square of error (RMSE) is 0.0018. Fitting of RMS by exponential function is depicted in Fig. 14, the R -Square of this fitting is 0.77, and its root mean square of error (RMSE) is 0.0016, so exponential function has a better fitting rather than Weibull basis hazard function and it is used for fitting of features. Fittings of absolute average value by Weibull basis hazard function and exponential

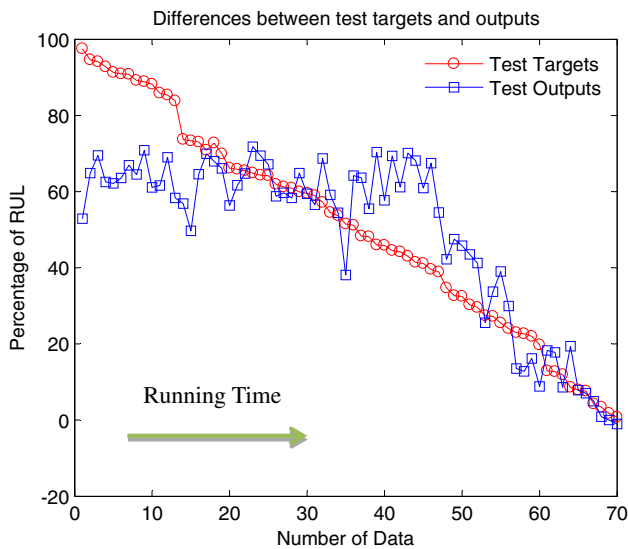


Fig. 10 Result of MLP training for data test 5

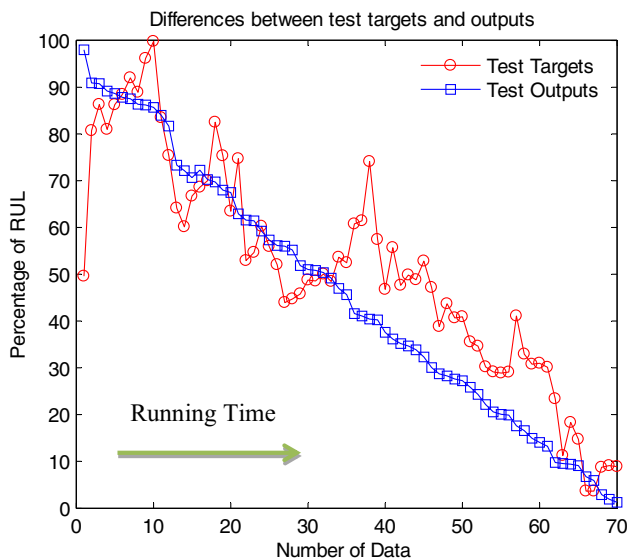


Fig. 12 Result of MLP with smoothed data

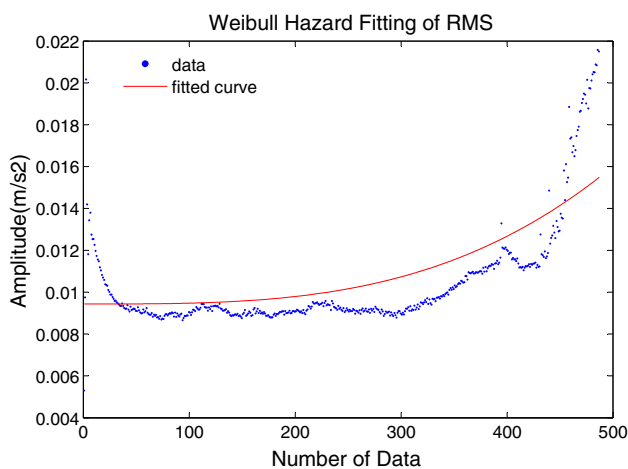


Fig. 13 Weibull basis hazard fitting of RMS

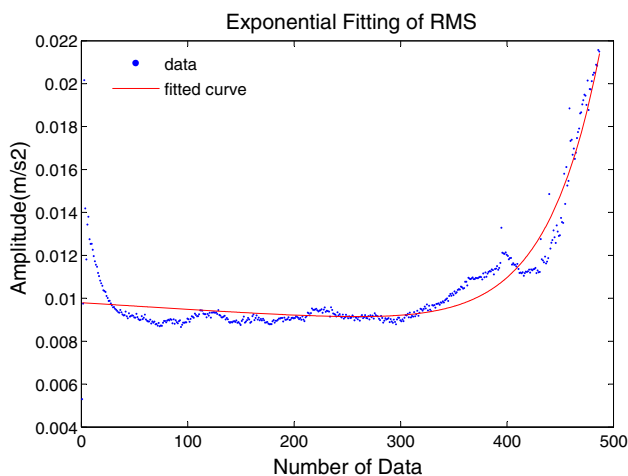


Fig. 14 Exponential fitting of RMS

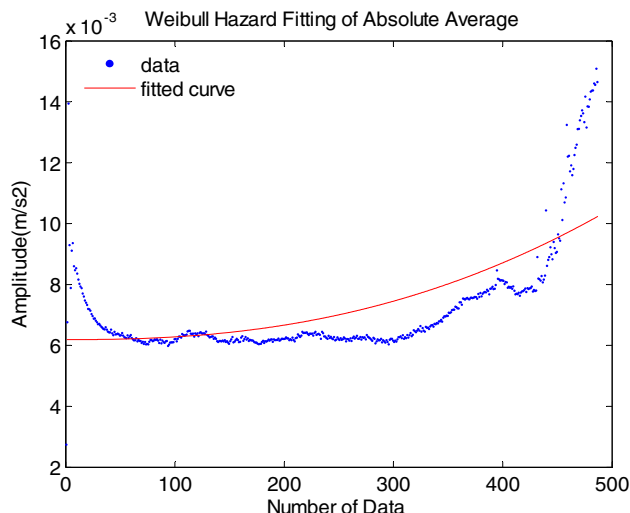


Fig. 15 Weibull basis hazard fitting of absolute average

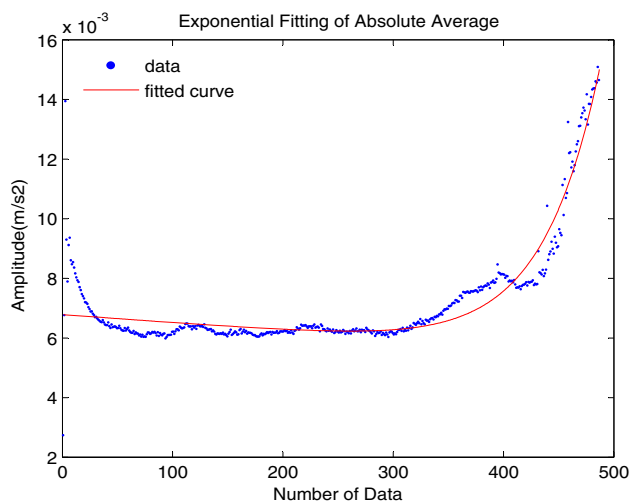


Fig. 16 Exponential fitting of absolute average

function are also depicted in Figs. 15 and 16. In this figure, exponential fitting also shows better performance than Weibull fitting.

The result of MLP network with smoothed and fitted data is depicted in Fig. 17. As shown in Fig. 17 and according to Table 3, the root mean square of error (RMSE) is 2.13 and there is a little error between targets and outputs.

Error of MLP network with raw input and preprocessed data is shown in Table 4. It can be seen from Table 4 that, as the preprocessing on the data is applied, accuracy of the network improves.

In the next step, the trained NN with data test 5 is used for predicting the RUL of data test 1 and data test 6. The training data for prediction of test 1 and test 6 contain data of test 5 and first half of the data of test 1 and first half of

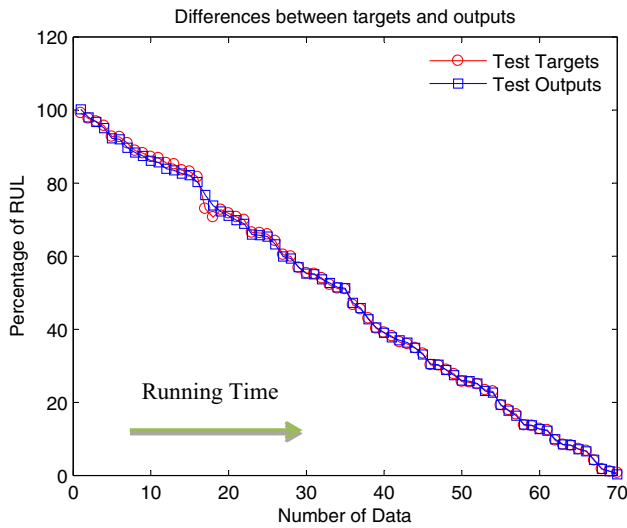


Fig. 17 Result of MLP with smoothed and exponential fitted data of test 5

Table 4 Comparison of error in MLP with different preprocessing data

Type of neural network	Root mean square of error, RMSE (%)
MLP with raw data from decomposed signal	15.83
MLP with smoothed data	11.46
MLP with smoothed and fitted data	2.13

the data of test 6, respectively. This ensures that the proposed prognostic method can be used on similar machines after training of the network with the data of only one machine. However, the prediction can be improved if the data of more than one machine are used for training.

The result of MLP network prediction with smoothed and fitted data of data test 1 is depicted in Fig. 18, and the result of MLP network prediction with smoothed and fitted data of data test 6 is depicted in Fig. 19.

As shown in Figs. 18 and 19, the predicted RUL is approaching to the desired or actual RUL and the network predicts the last time of experiment with a small error. The error of MLP neural network with smoothed and fitted data for data test 1 and data test 6 is shown in Table 5. As shown in Table 5, there is little error between actual and predicted results.

It is interesting to compare the results of the current method with other references. Because error is reported in various forms in different papers, it is not possible to compare the results directly. A descriptive comparison is given as follows:

In Gebraeel and Lawley (2008), the error in RUL prediction is about 7.56% and other benchmarks pointed to in

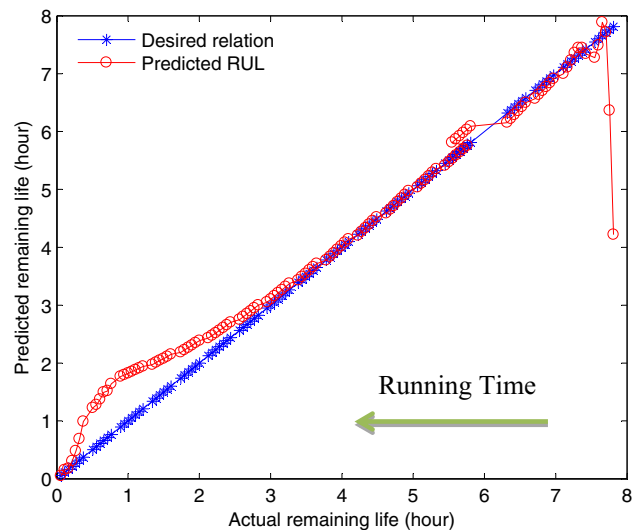


Fig. 18 Relation between predicted RUL and actual RUL for data test 1 (only first 50% of the data is used for training)

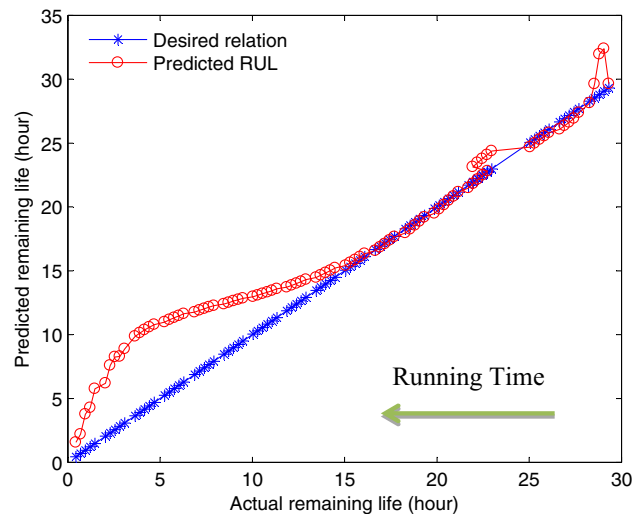


Fig. 19 Relation between predicted RUL and actual RUL for data test 6 (only first 50% of the data is used for training)

Table 5 Comparison of error in MLP for other data tests

Type of NN, test no.	Root mean square of error, RMSE (h)
MLP with smoothed and fitted data, data test 1	0.48
MLP with smoothed and fitted data, data test 6	2.92

that paper have a high error compared to current proposed method.

In Huang et al. (2007), errors were grouped into some classes (less than 10%, greater than 10% but not greater

than 20% and greater than 20%), and generally error of their method is high.

In Kim et al. (2010), predicted RUL is based on health state and error of their method in the first part of the test is high. Error value decreases at the end of the test.

In Sloukia et al. (2013), mixture of hidden Markov model and support vector machine has been used and they reported relatively high error between estimated and real RUL (38.52 and 21.85%).

In Gebraeel et al. (2009), predicted RUL of bearing by using Bernstein fitting and despite the complexity of computation, the average prediction error is reported between 16.2 and 22.5%.

As can be seen, the proposed method has less error compared to the published results.

6 Conclusions

This paper is intended to present an efficient and accurate method for RUL prediction of REBs. In this work, vibration signal is decomposed by WPT. New methods for selecting appropriate mother wavelet, optimal level and optimal node are proposed. The selected features are extracted from the decomposed signal and then smoothed and fitted by Weibull hazard function as well as exponential function. Effect of preprocessing of the input data on the prediction results of the MLP network was investigated, and it was concluded that MLP with smoothed and fitted data has better RUL prediction than other methods. Additionally it was verified that the trained network for one test can be applied in prediction of other data tests successfully. In this study, it has been assumed that rotation speed and load are constant. If the operating conditions of the machine are not constant, the current approach shall be modified to account for variation in degradation rate of bearing. This case which can be observed in some applications is the subject of a new research. The proposed approach can be considered as the first investigative step since it concerns a single application of the method to specific type of machines and to unique specimens and therefore its effectiveness for other components and various cases has to be proved with further investigations.

Acknowledgements Experimental tests of this research were conducted in the vibration laboratory of Sharif University of Technology. Hereby, authors express their thanks to Professor Mehdi Behzad for his kind support and guidance.

References

Akbari M, Homaei H, Heidari M (2014) An intelligent fault diagnosis approach for gears and bearings based on wavelet transform as a

preprocessor and artificial neural networks. *Int J Math Model Comput* 4(4):309–329

Al-Badour F, Sunar M, Cheded L (2011) Vibration analysis of rotating machinery using time–frequency analysis and wavelet techniques. *Mech Syst Signal Process* 25(6):2083–2101

Ali JB, Chebel-Morello B, Saidi L, Malinowski S, Fnaiech F (2015) Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network. *Mech Syst Signal Process* 56:150–172

Antoni J (2009) Cyclostationarity by examples. *Mech Syst Signal Process* 23(4):987–1036

Bafroui HH, Ohadi A (2014) Application of wavelet energy and Shannon entropy for feature extraction in gearbox fault detection under varying speed conditions. *Neurocomputing* 133:437–445

Chen X, Shen Z, He Z, Sun C, Liu Z (2013) Remaining life prognostics of rolling bearing based on relative features and multivariable support vector machine. *Proc Inst Mech Eng Part C J Mech Eng Sci* 227(12):2849–2860

El-Thalji I, Jantunen E (2015) A summary of fault modelling and predictive health monitoring of rolling element bearings. *Mech Syst Signal Process* 60:252–272

Gebraeel NZ, Lawley MA (2008) A neural network degradation model for computing and updating residual life distributions. *IEEE Trans Autom Sci Eng* 5(1):154–163

Gebraeel N, Pan J (2008) Prognostic degradation models for computing and updating residual life distributions in a time-varying environment. *IEEE Trans Reliab* 57(4):539–550

Gebraeel N, Lawley M, Liu R, Parmeshwaran V (2004) Residual life predictions from vibration-based degradation signals: a neural network approach. *IEEE Trans Ind Electron* 51(3):694–700

Gebraeel N, Lawley M, Li R, Ryan J (2005) Residual-life distributions from component degradation signals: a Bayesian approach. *IIE Trans* 37(6):543–557

Gebraeel N, Elwany A, Pan J (2009) Residual life predictions in the absence of prior degradation knowledge. *IEEE Trans Reliab* 58:106–116

Goebel K, Eklund N, Bonanni P (2006) Fusing competing prediction algorithms for prognostics. In: 2006 IEEE Aerospace Conference. IEEE, pp 10–pp

Guo L, Li N, Jia F, Lei Y, Lin J (2017) A recurrent neural network based health indicator for remaining useful life prediction of bearings. *Neurocomputing* 240:98–109

Harris TA (2001) *Rolling bearing analysis*. Wiley, New York

Heng A, Zhang S, Tan AC, Mathew J (2009) Rotating machinery prognostics: state of the art, challenges and opportunities. *Mech Syst Signal Process* 23(3):724–739

Hong S, Zhou Z, Zio E, Hong K (2014) Condition assessment for the performance degradation of bearing based on a combinatorial feature extraction method. *Digit Signal Process* 27:159–166

Howard I (1994) Review of rolling element bearing vibration detection, diagnosis and prognosis (No. DSTO-RR-0013). Defence Science and Technology Organization Canberra (Australia)

Hu C, Youn BD, Wang P, Yoon JT (2012) Ensemble of data-driven prognostic algorithms for robust prediction of remaining useful life. *Reliab Eng Syst Saf* 103:120–135

Huang R, Xi L, Li X, Liu CR, Qiu H, Lee J (2007) Residual life predictions for ball bearings based on self-organizing map and back propagation neural network methods. *Mech Syst Signal Process* 21(1):193–207

Jammu NS, Kankar PK (2011) A review on prognosis of rolling element bearings. *Int J Eng Sci Technol* 3(10):7497–7503

Jardine AKS, Makis V, Banjevic D, Braticevic D, Ennis M (1998) A decision optimization model for condition-based maintenance. *J Qual Maint Eng* 4(2):115–121

- Kankar PK, Sharma SC, Harsha SP (2011) Rolling element bearing fault diagnosis using wavelet transform. *Neurocomputing* 74(10):1638–1645
- Kim HE, Tan AC, Mathew J, Kim EY, Choi BK (2010) Prognosis of bearing failure based on health state estimation. In: Kiritisis D, Emmanouilidis C, Koronios A, Mathew J (eds) *Engineering asset lifecycle management*. Springer, London, pp 603–613
- Kim HE, Tan AC, Mathew J, Choi BK (2012) Bearing fault prognosis based on health state probability estimation. *Expert Syst Appl* 39(5):5200–5213
- Kotzalas MN, Harris TA (2001) Fatigue failure progression in ball bearings. *Trans Am Soc Mech Eng J Tribol* 123(2):238–242
- Kulkarni PG, Sahasrabudhe AD (2013) Application of wavelet transform for fault diagnosis of rolling element bearings. *Int J Technol Enhanc Emerg Eng Res* 2(4):138–148
- Kumar HS, Srinivasa Pai P, Sriram NS, Vijay GS (2014) Selection of mother wavelet for effective wavelet transform of bearing vibration signals. *Adv Mater Res* 1039:169–176
- Li Y, Billington S, Zhang C, Kurfess T, Danyluk S, Liang S (1999) Adaptive prognostics for rolling element bearing condition. *Mech Syst Signal Process* 13(1):103–113
- Li Y, Kurfess TR, Liang SY (2000) Stochastic prognostics for rolling element bearings. *Mech Syst Signal Process* 14(5):747–762
- Liu J, Wang W, Golnaraghi F (2009) A multi-step predictor with a variable input pattern for system state forecasting. *Mech Syst Signal Process* 23(5):1586–1599
- Liu Z, Zuo MJ, Qin Y (2016) Remaining useful life prediction of rolling element bearings based on health state assessment. *Proc Inst Mech Eng Part C J Mech Eng Sci* 230(2):314–330
- Loutas T, Roulias D, Georgoulas G (2013) Remaining useful life estimation in rolling bearings utilizing data-driven probabilistic e-support vectors regression. *IEEE Trans Reliab* 62(4):821–832
- Marble S, Morton BP (2006) Predicting the remaining life of propulsion system bearings. In: 2006 IEEE Aerospace conference. IEEE, pp 8–pp
- Nandi S, Toliyat HA, Li X (2005) Condition monitoring and fault diagnosis of electrical motors—a review. *IEEE Trans Energy Convers* 20(4):719–729
- Nikolaou NG, Antoniadis IA (2002) Rolling element bearing fault diagnosis using wavelet packets. *Ndt E Int* 35(3):197–205
- Peng ZK, Chu FL (2004) Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography. *Mech Syst Signal Process* 18(2):199–221
- Peng Y, Cheng J, Liu Y, Li X, Peng Z (2018) An adaptive data-driven method for accurate prediction of remaining useful life of rolling bearings. *Front Mech Eng* 13(2):301–310
- Pham HT, Yang BS (2010) Estimation and forecasting of machine health condition using ARMA/GARCH model. *Mech Syst Signal Process* 24(2):546–558
- Qian Y, Yan R, Gao R (2017) A multi-time scale approach to remaining useful life prediction in rolling bearing. *Mech Syst Signal Process* 83:549–567
- Qiu H, Lee J, Lin J, Yu G (2003) Robust performance degradation assessment methods for enhanced rolling element bearing prognostics. *Adv Eng Inf* 17(3):127–140
- Rafiee J, Tse PW, Harifi A, Sadeghi MH (2009) A novel technique for selecting mother wavelet function using an intelligent fault diagnosis system. *Expert Syst Appl* 36(3):4862–4875
- Rai A, Upadhyay S (2018) Intelligent bearing performance degradation assessment and remaining useful life prediction based on self-organising map and support vector regression. *Proc Inst Mech Eng Part C J Mech Eng Sci* 232(6):1118–1132
- Rao BKN, Pai PS, Nagabhushana TN (2012) Failure diagnosis and prognosis of rolling-element bearings using artificial neural networks: a critical overview. In: *Journal of Physics: Conference Series*, vol 364, No. 1). IOP Publishing, p 012023
- Reuben LCK, Mba D (2014) Bearing time-to-failure estimation using spectral analysis features. *Struct Health Monit* 13(2):219–230
- Shao Y, Nezu K (2000) Prognosis of remaining bearing life using neural networks. *Proc Inst Mech Eng Part I J Syst Control Eng* 214(3):217–230
- Si XS, Wang W, Hu CH, Zhou DH (2011) Remaining useful life estimation—a review on the statistical data driven approaches. *Eur J Oper Res* 213(1):1–14
- Singh S, Howard CQ, Hansen CH (2015) An extensive review of vibration modelling of rolling element bearings with localised and extended defects. *J Sound Vib* 357:300–330
- Sloukia F, El Aroussi M, Medromi H, Wahbi M (2013) 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, pp 1–4
- Tian Z, Wong L, Safaei N (2010) A neural network approach for remaining useful life prediction utilizing both failure and suspension histories. *Mech Syst Signal Process* 24(5):1542–1555
- Wang W (2007) An adaptive predictor for dynamic system forecasting. *Mech Syst Signal Process* 21(2):809–823
- Wang D, Tsui KL (2017) Statistical modeling of bearing degradation signals. *IEEE Trans Reliab* 66(4):1331–1344
- Wang Y, Peng Y, Zi Y, Jin X, Tsui KL (2015) An integrated Bayesian approach to prognostics of the remaining useful life and its application on bearing degradation problem. In: 2015 IEEE 13th international conference on industrial informatics (INDIN). IEEE, pp 1090–1095
- Yan R (2007) Base wavelet selection criteria for non-stationary vibration analysis in bearing health diagnosis, Doctoral Dissertations Available from Proquest. AAI3275786
- Yiakopoulos CT, Antoniadis IA (2005) Cyclic bispectrum patterns of defective rolling element bearing vibration response. *Forsch Ingenieurwes* 70(2):90–104
- Zhang Z, Wang Y, Wang K (2013) Fault diagnosis and prognosis using wavelet packet decomposition, Fourier transform and artificial neural network. *J Intell Manuf* 24(6):1213–1227
- Zhao M, Tang B, Tan Q (2016) Bearing remaining useful life estimation based on time–frequency representation and supervised dimensionality reduction. *Measurement* 86:41–55