2017 International Conference on Sensing, Diagnostics, Prognostics, and Control

Remaining Useful Life Prediction of Bearings Using Fuzzy Multimodal Extreme Learning Regression

Xuejiao Li
Department of Architecture and Civil Engineering
Chongqing Telecommunication Polytechnic College
Chongqing 402247, China
1044109363@qq.com

***Abstract*—Remaining Useful Life (RUL) prediction of bearings is one of the crucial conditions for timely maintenance. In this paper, a fuzzy multimodal extreme learning regression is proposed for the RUL estimation. In this method, fuzzy fusion, ensemble empirical mode decomposition (EEMD), and extreme learning machine (ELM) are integrated. The fuzzy fusion is first used to fuse original features for establishing a condition criterion. EEMD is subsequently utilized to decompose the condition criterion into several sub-series of multiple modes. EML is then adopted for predicting the sub-series in each mode. The predicted sub-results in each mode are finally summarized as the final results. The proposed method is assessed by the bearings data from NSF I/UCR center. Experimental results reveal that the proposed approach is able to build the condition criterion to reflect bearings degradation. It performs better in the RUL prediction than benchmark approaches.**

Keywords—bearings; remaining useful life; fuzzy fusion; ensemble empirical mode decomposition; extreme learning machine

I. INTRODUCTION

In recent years, many researchers have paid much attention to prognostics and health management (PHM) of bearings as the ability to reduce the maintenance cost by estimating remaining useful life (RUL) while increasing the reliability and safety of bearings. Bearings are the critical elements widely used in rotating machinery. Different kinds of faults have occurred as bearings work in the tough environments as the result the machine's operation will be affected. In the various faults of rotating machinery, approximately 30% failures are caused by the rolling bearings [1].

Contemporary, the mounting number of the prediction models has been proposed. Nevertheless, the most daunting trouble is how to identify a bearing's fault before it reaches a critical level. Thus it is significant to predict the RUL of bearings before the damage occurs.

The RUL prediction of bearings is not only a challenging subject but also a hot issue. In recent years, two basic prognostic models are commonly used to predict RUL. One is model-based method. It relies on empirical or physics-based models to describe degradation processes and adjust model parameters according to the information from measurements [2]. Another is data-driven method. It is able to forecast the future trend of components by taking full advantage of the historical data, such as artificial neural network [3], support vector regression (SVR) [4], neuro-fuzzy system [5], and particle filtering [6]. Allowing for the fast learning capacity, extreme learning machine (ELM) is adopted in this work.

Vibration signal analysis is one of the most effective methods for the bearings condition monitoring [7]. According to the data-driven method, the RUL prediction of bearings is composed of two steps: firstly, a criterion reflecting the degradation of bearings should be constructed. Secondly, an appropriate regression model is employed for predicting the RUL of bearings [8]. It is very important to erect a suitable criterion in the prediction process. The features are extracted from time domain and time-frequency domain which have been used to construct the condition criterion. Nonetheless it is not reliable to reflect the bearing degradation processes if only inadequate features are used. So many approaches are developed to fuse multiple features of bearings for building a condition criterion. In the fusion process, there may be a redundancy between the features. To address this problem, fuzzy fusion approach is adopted in this paper. Additionally, ensemble empirical mode decomposition (EEMD) is introduced to improve the performance of the regression model. Given the above mentioned, fuzzy multimodal extreme learning regression (FMELR) is proposed to predict the RUL of bearings in this work.

II. METHODOLOGY

A. Fuzzy Fusion

Aiming at the multiple characteristic indexes, a fusion data method is adopted. The cost function is expressed as follow:

$$G_n = F(C_u(a_n(v))) \, , \tag{1}$$

where $F(.)$ means data fusion function, C_u represents the uth indicator.

Next, using the fuzzy progress to achieve the data fusion function $F(.)$. The set of initial minimum spectral subset is $X(f) = \{A_1, A_2, \ldots, A_m, \ldots, A_n\}$, and the time domain component set is $\{a_1(v), a_2(v), \ldots, a_m(v), \ldots, a_n(v)\}$.

Then each index is normalized between [0,1],

978-1-5090-4020-9/17 $31.00 © 2017 IEEE

DOI 10.1109/SDPC.2017.100

$$MC_u = \{MA_1, MA_2, ..MA_m, ..MA_n\}\ , \tag{2}$$

Calculate the mean $\overline{MC_u}$ and standard deviation σ_u of each normalized index sequence,

$$\overline{MC_u} = \frac{1}{n}\sum_{m=1}^{n} MC_u\ , \tag{3}$$

$$\sigma_u = \frac{1}{n-1}\sqrt{\sum_{m=1}^{n}(MC_u \quad \overline{MC_u})^2}\ . \tag{4}$$

Then further calculate the grand average and the total standard deviation of the n index sequence. In this paper, n represents used indicators.

The triangular fuzzy membership function is chosen, and the fuzzy quantity of each index is given as:

$$\widetilde{C}_U = (q_{u1}, q_{u2}, q_{u3}) = (\overline{MC_u} \quad 2\sigma_u, \overline{MC_u}, \overline{MC_u} + 2\sigma_u)\ . \tag{5}$$

Similarly, the fuzzy quantity of the cost function of data fusion is expressed as:

$$\widetilde{C}_0 = (q_{01}, q_{02}, q_{03}) = (\overline{MC_0} \quad 2\sigma_0, \overline{MC_0}, \overline{MC_0} + 2\sigma_0)\ . \tag{6}$$

The relationship between the fuzzy quantity of the u indexes and the data fusion fuzzy quantity is:

$$P_u = \frac{1}{1 + \frac{q_{u1} + 4q_{u2} + q_{u3} \quad q_{01} \quad 4q_{02} \quad q_{03}}{6}}\ . \tag{7}$$

The bigger the Pu is, the closer the fuzzy quantity of the u index and the fuzzy quantity of data fusion are. Otherwise, they will be inconformity. Therefore, the fuzzy quantity of a certain index is closer to the data fusion fuzzy quantity, the better its reliability and stability are, and its weight should be larger [9].

B. Ensemble Empirical Mode Decompositon

EEMD [10] as an improved version of empirical mode decomposition (EMD), adsd the white noise series into the original signal provides a uniform reference scale distribution to promote the EMD and establish a reference frame in the time-frequency space, which is proposed to overcome the mode mixing drawback produced by the EMD. The implementation procedure of the EEMD method can be found in these literatures. Brief description is as follows.

Add a random white noise signal d_z to original signal w,

$$W_z = w + d_z\ , \tag{8}$$

where W_z is the noise-added signal, z=1, 2,…, N, and N is the trial number.

Decompose the signal X_t which added white nose into IMFs,

$$W_z = \sum_{t=1}^{N} C_{t,z} + r_{z,N+1}\ , \tag{9}$$

where $C_{t,z}$ represents the tth IMF of the zth trial, $r_{o,N+1}$ is the residue of zth trial, and N+1 is the number of IMFs.

Compute the ensemble mean c_t of the ensemble members R for each IMF,

$$c_t = \frac{1}{R}\sum_{z=1}^{R} C_{t,z}\ . \tag{10}$$

The added white noise and the effect are obtained by the following equation:

$$e = \frac{\omega}{\sqrt{R}}\ , \tag{11}$$

where R means ensemble members, ω is amplitude of the added noise, and e is the standard deviation of error. According to the previous study, in this paper, R=500 and ω=0.2.

C. Extreme Learning Machine

ELM is a kind of effective learning algorithms for classification and prediction in single-hidden layer feed-forward neural network (SLFNs) [11]. Fig.1 shows the structure of the ELM model.

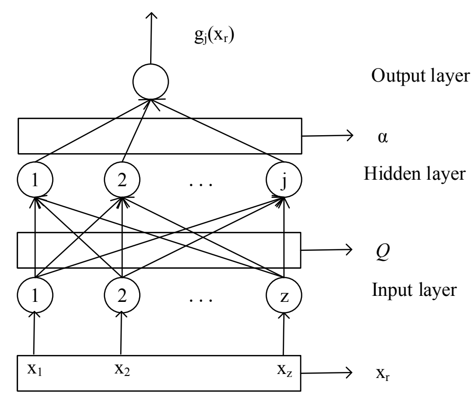


Figure 1. The structure of ELM model

Suppose that one train SLFNs with Q hidden neurons and activation function $f(x)$ to learn K distinct samples (c_i, b_i), where $c_i=[c_{i1}, c_{i2}, \ldots, c_{is}]^T \in R^s$ denotes the feature and $b_i=[b_{i1}, b_{i2}, \ldots, b_{ih}] \in R^h$ denotes the target value. The SLFNs with $j \leq K$ hidden nodes, and is mathematically described as:

$$g_j(x_r) = \sum_{i=1}^{j} \alpha_i f(c_i, b_i, x_r) = S_r\ , \tag{12}$$

where r=1, 2, …, K, i=1,…, j, α_i connects the ith hidden node and the output node, and S_r is the predicted output value. If an ELM can approximate these j samples with zero error, then the equation can be represented as:

$$\sum_{i=1}^{j} \alpha_i f(c_i, b_i, x_r) = y_r\ ,\ r=1,\ 2,\ \ldots,\ K. \tag{13}$$

The above equations can be written compactly as:

$$D\alpha = B\ , \tag{14}$$

$$D = \begin{bmatrix} g(c_1, b_1, x_1) & \cdots & g(c_j, b_j, x_1) \\ \vdots & \cdots & \vdots \\ g(c_1, b_1, x_r) & \cdots & g(c_j, b_j, x_K) \end{bmatrix}_{K \times j}, \qquad (15)$$

where D is the hidden layer output neurons, α is the matrix of output weights and B is the matrix of targets.

In ELM, the output weight α can be written as:

$$\alpha = D^{+} B, \qquad (16)$$

$$D^{+} = (D^T D)^{-1} D^T. \qquad (17)$$

Although ELM learning speed is faster than other machine learning algorithms, its accuracy still depends on feature extraction [12].

D. The FMELR Model

Having presented the constituents separately, steps of the methodology application of the FMELR prediction approach is illustrated in Fig. 2.

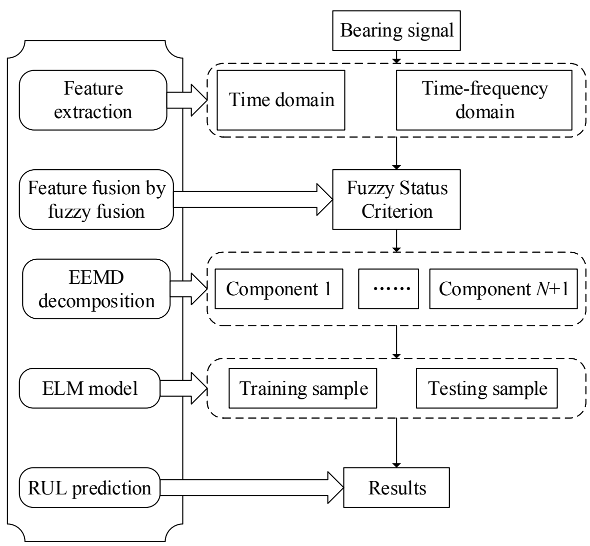


Figure 2. Procedure of FMELR

Step 1: Collect vibration signal of bearing, and extract the features from multiple domains.

Step 2: Apply the fuzzy method to fuse the multi-features as a status criterion.

Step 3: Employ the EEMD decompose the criterion into N modes.

Step 4: Build the ELM models for the corresponded components.

Step 5: The prediction sub-results in different modes are added correspondingly into the final prediction result.

III. Application and Analysis

In this section, the proposed FMELR is assessed by predicting RUL of bearing. In order to verify the effectiveness of the method, accelerated degradation tests of bearing are conducted and vibration signals are acquired during the tests. Moreover, the peer models are performed for validating the proposed model.

A. Experimental Data

The experimental data comes from the NSF I/UCR Center for Intelligent Maintenance Systems (IMS). During the test, double row bearing is installed on the shaft. The rotation speed was kept constantly at 2000 r/min and 6000 1b radial load is placed onto the shaft and bearing by a spring mechanism. The sampling rate is 20 kHz and the data length includes 20480 points. Vibration signals are collected every 10 minutes by a NI DAQ Card 6062E. At the end of the run-to-failure experiment a crack is found near the shoulder of bearing [13]. The raw vibration signals is shown in Fig. 3

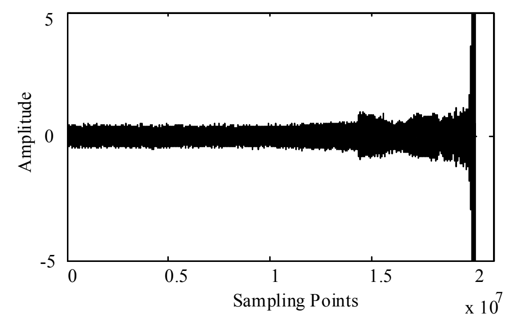


Figure 3. Vibration signals of bearing

B. Status Criterion of Bearing Establishment

Feature extraction is an important step for the RUL prediction of bearings. The features are extracted from time-domain, and time-frequency domain of the raw vibration signals including standard deviation, variance, kurtosis, maximum, peak-to peak value, absolute mean value, root mean square, and wave form factor, which are employed from the time domain. Furthermore, wavelet transform package is used to obtain 8 features extracted from time-frequency domain. And the energy values of 8 frequency bands decomposed by the wavelet packet of the three layers are used as time-frequency characteristics. Totally, there are 16 features for establishing the fusion status criterion. Furthermore the fuzzy fusion method is used to carry on the fusion of the 16 features into one status criterion, shown in Fig.4.

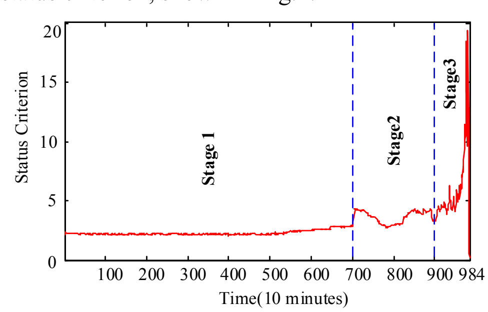


Figure 4. Status criterion for bearing

Fig. 4 shows the shape of vibration signals more clearly after a fuzzy fusion algorithm based on the data. In the whole lifetime of a bearing, it is obvious that the process of degradation has three distinct stages including the normal operation stage, the slight degradation stage, and the serious degradation stage. The state criterion keeps stable during the stage1 and increase rapidly from stage 2 to stage 3 until failure. It also shows that the extracted state criterion is effective and can significantly reflect the life changing process. The stages 2 and 3 (700-984) of the status criterion are used for modeling on account of bit information in stage 1.

C. EEMD Decomposition

Apply EEMD to decompose of the status criterion after use fuzzy fusion. The decomposition results are given in Fig. 5.

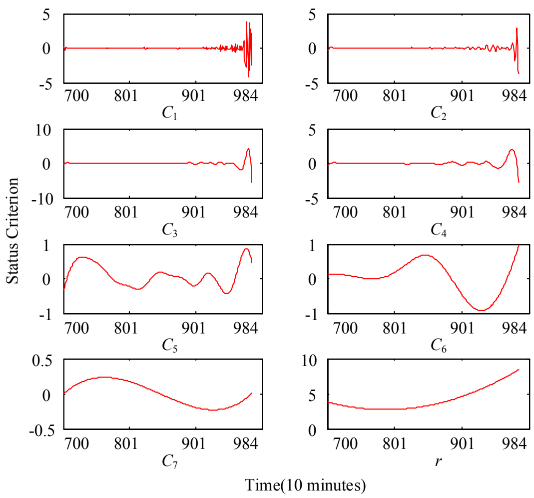


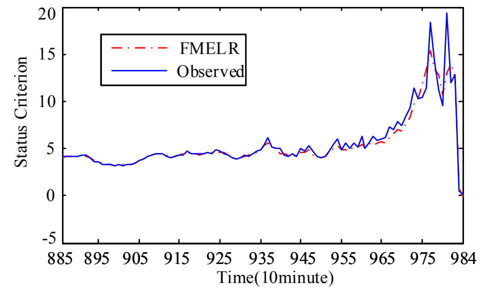
Figure 5. Decompositon of the status criterion

The status criterion is decomposed into seven independent components C_1-C_7, and one residue component r. They describes the holistic trend of status criterion as a monotonic function. After decomposing the status criterion , the original data becomes more stable and regular.

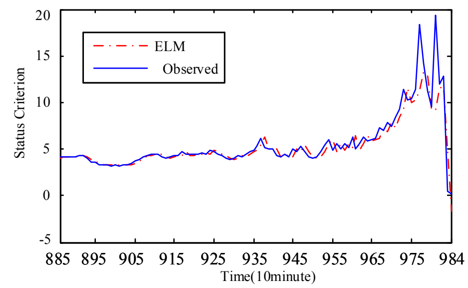
EEMD weakens the coupling and interruption to some extent among the information that contains the raw data. According to the proposed method, the eight components are used to build eight ELM models respectively.

D. RUL Prediction

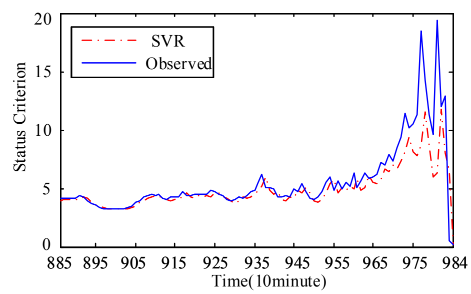
ELM input parameters include the eight components to predict the RUL in different modes. Before data input, training sample andtest sample sets are established for this model. In this work, 65% of the sample sets are used for training and the rest of are used for testing. Mean absolute error (MAE), root mean square error (RMSE) and correlation coefficient (CC) indicators are used as the criteria to evaluate the performance of the model. To highlight the performance of the proposed model, the standard ELM is employed to compare the effect of the multi-mode prediction. In many fields, SVR is adopted, because it is not only anuseful but also a commonly used tool for complex models and nonlinear systems. The same dataset used tried to compare the prediction results in Fig.6.



(a)



(b)



(c)

Figure 6. The prediction results for different models:
(a) FMELR; (b) ELM; (c) SVR

Figs.6 (a)-(c) represents the changing trend of FMELR, and ELM prediction values are close to the observed values. Because the fitting effect is superior so that is utilized to predict extreme points.

Results obtained by FMELR are much closer to the true values than ELM. At the same time, compared to the other two models, the effect of RUL prediction on SVR is more worse. It captures the trend of observed values at the beginning. However, in the part of violent fluctuation, it is difficult to grasp the laws of SVR and it is the worst prediction results for the extreme points among three models. The rationality of the proposed model is demonstrated by the comparative analysis of three models between experimental and theoretical results. As shown in Fig. 7, compared with SVR and ELM, FMELR has the lowest absolute mean error with the best stability. By observing the length of each box entity, results generated by FMELR and ELM are shorter than those of SVR.All of these indicate the absolute error distributions of these models are relatively concentrated. The location of the entity for FMELR is lower than SVR, hence, most of absolute error of FMELR is within lower range. Through comparison between median and quartiles about distance, the situation of FMELR is relatively symmetrical and with a basically normal distribution. In addition, all the peers are not balanced obviously. This shows that FMELR is good in terms of the distribution of absolute error. It is noteworthy that all models have some outliers with but SVR has moreoutliers than other two. This also proves the advantage of ELM. Through the above analyses, the superiority of the present FMELR model comparing to its peers is presented.

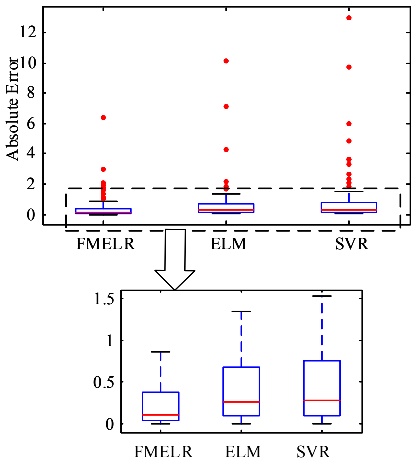


Figure 7. Comparisons of boxplot of absolute errors using different models

Table I shows that the FMELR model shows better accuracy and performance than the ELM and the SVR in terms of the various performance criteria. MAE and RMSE and CC of FMELR model is the lowest than the other two models, These differences also illustrate that multi-features fuzzy fusion and multimodal decomposition can better extract essential features of the processing of degradation.

TABLE I. COMPARION OF MODELS

Model	Criteria		
	MAE	***RMSE***	***CC***
SVR	0.861	2.023	0.813
ELM	0.620	1.444	0.890
FMELR	0.381	0.875	0.960

According to the above comprehensive analysis and comparison, the FMELR using multimodal features learning has better prediction performance than the ELM and the SVR, which can satisfy the requirements for the RUL of bearing.

IV. CONCLUSION

In this work, an approach called FMELR combining the EEMD and the ELM is proposed for predicting the RUL of bearing. In this method, the features from two domains of the bearing are extracted to fuse a status criterion which can reflect the RUL of bearing. The EEMD is employed to decompose the status criterion into a multimodal subseries. For each model, the ELM is used for subseries regression. Adding up the sub-prediction results become the final prediction results.

The proposed FMELR approach is evaluated in the full set data of bearing from the NSF I/UCR Center on Intelligent Maintenance Systems. Besides, the ELM and the SVR are used as reference models. Experimental results indicate that the proposed FMELR method has better prediction capacity than its peers in terms of MAE, RMSE, and CC. The proposed method generates more reliable and accurate prediction results, attributes to multi-features fuzzy fusion, multimodal information mining, and high-quality features learning.

ACKNOWLEDGMENT

This work is supported in part by the National Key Research & Development Program of China (2016YFE0132200), and the Chongqing Science & Technology Commission (cstc2015zdcy-ztzx70012).

REFERENCES

[1] Q. He, L. Li, G. Jiang, and P. Xie, “Residual life predictions for bearings based on PCA and MELM,” China Mechanical Engineering, pp.984-989, 2014.

[2] Y. Lei, N. Li, F. Jia, and J. Lin, “A nonlinear degradation model based method for remaining useful life prediction of rolling element bearings,” Prognostics and System Health Management Conference, pp. 1-8, 2015.

[3] A. K. Mahamad, S. Saon and T. Hiyama, “Predicting remaining useful life of rotating machinery based artificial neural network,” Computers and Mathematics with Applications, vol. 60, pp. 1078-1087, 2010.

[4] H. Z. Huang, H. K. Wang, Y. F. Li, L. Zhang, and Z. Liu, “Support vector machine based estimation of remaining useful life: current research status and future trends,” J. Mech. Sci. Technol., vol. 29, pp.151-163, 2015.

[5] C. Chen, G. Vachtsevanos and M. E. Orchard, “Machine remaining useful life prediction: an integrated adaptive neuro-fuzzy and high-order particle filtering approach,” Mech. Syst. Signal Pr, vol. 28, pp. 597-607.

[6] E. Zio and G. Peloni, “Particle filtering prognostic estimation of the remaining useful life of nonlinear components,” Reliability Engnieering and System Safety, vol. 96, pp. 403-409, 2011.

[7] C. Li, D. Cabrera, JVD. Oliveira, RV. Sanchez, M. Cerrada, and G. Zurita, “Extracting repetitive transients for rotating machinery diagnosis using multiscale clustered grey infogram,” Mech. Syst. Signal Pr., vol. 76, pp. 157-173, 2016.

[8] X. Chen, Z. Shen, Z. He, C. Sun, and Z. Liu, “Remaining life predictions of rolling bearing based on relative features and multivariable support vector machine,” J. Mech. Eng-en., vol. 227, pp. 2849-2860, 2013.

[9] C. Li, M. Liang and T. Wang “Criterion fusion for spectral segmentation and its application to optimal demodulation of bearing vibration signals,” Mech. Syst. Signal Pr., vol. 64, pp. 132-148, 2015.

[10] Z. H. Wu and N. E. Huang, “Ensemble empirical mode decompositon: a noise-assisted data analysis mothod,” Advances in Adaptive Date Analysis, pp. 1-41, 2011.

[11] Z. L. Sun, T. M. Choi, K. F.Au, and Y. Yu, “Sales forecasting using extreme learning machine with application in fashion retailing,” Decision Support Systems, vol. 46, pp. 411-419, 2009.

[12] Y. Liu, B. He, F.Liu, S. Lu, Y. Zhao,and J. Zhao, “Remaining useful lLife prediction of rolling bearings using PSR, JADE, and extreme learning machine,” Math. Probl. Eng., pp. 1-13, 2016.

[13] H. Qiu, J. Lee and J. Lin, “Wavelet filter-based weak signature detection method and its application on roller bearing prognostics,” J. Sound Vib., vol. 289, pp. 1066–1090, Feb 2006.