MATEC Web of Conferences **255**, 02003 (2019) https://doi.org/10.1051/matecconf/201925502003
EAAI Conference 2018

Optimized ELM based on Whale Optimization Algorithm for gearbox diagnosis

M. Firdaus Isham[1,2*], *M. Salman* Leong[1], *M. H.* Lim[1] and *Z. A.B.* Ahmad[2]

[1] Institute of Noise and Vibration, Universiti Teknologi Malaysia, Jalan Sultan Yahya Petra, 54100 Kuala Lumpur, Malaysia

[2] Sekolah Kejuruteraan Mekanikal, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia

Abstract. Extreme learning machine (ELM) is a fast and quick learning algorithm with better generalization performance. However, the randomness of input weight and hidden layer bias may affect the overall performance of ELM. This paper proposed a new approach to determine the optimized values of input weight and hidden layer bias for ELM using whale optimization algorithm (WOA), which we call WOA-ELM. An online gearbox vibration signals is used in this study. Empirical mode decomposition (EMD) and complementary mode decomposition (CEEMD) are used to decompose the signals into sub-signals known as intrinsic mode functions (IMFs). Then, statistical features are extracted from selected IMFs. WOA-ELM is used for classification of healthy and faulty condition of gearbox. The result shows that WOA-ELM provide better classification result as compared with conventional ELM. Therefore, this study provide a new diagnosis approach for gearbox application.

1 Introduction

Gearbox diagnosis study is one of the most important research study as gearbox is a crucial engineering unit in many industry. The study also help to ensure the functionality and reliability of gearbox applications and avoiding failures [1,2]. Advancement of technology has make the diagnosis approach has been improved from time to time. Nowadays, diagnosis based on frequency spectrum is no longer efficient as it may lead to inaccurate result which can lead to machine breakdown and catastrophic failure [3,4].

Automated diagnosis approach is known as a recent trend for diagnosis study. Basically, it consists of signal processing, features extraction, features selection and machine learning algorithm. A lots of research using automated diagnosis approach has been proposed recently in many areas of study [5–8].

Extreme learning machine (ELM) is a technique to train single-hidden layer feedforward neural network (SLFN) that has been proposed by Huang, Zhu and Siew in 2006 [9]. ELM has gained a lots of attention from many researcher since then. It has been applied in many area of study such as rotating machinery application [10–12], air quality forecasting [13], electric load forecasting [14], machining [15], etc. By comparing with conventional neural network, ELM provide thousands of time faster learning, better generalization and simple model [9]. Basically, ELM assign the input weight and hidden layer bias randomly. This make the algorithm less effective which subject to the given values [16]. Also, it also effect the ELM performance and provide unstable ELM output [17].

This paper aim to propose an approach to determine the optimized value of input weight and hidden layer bias by using whale optimization algorithm (WOA), so called as WOA-ELM. WOA is a new meta-heuristic method proposed by Mirjalili and Lewis in 2016 inspired by whale feeding behaviour of humpback whales [18]. According to Mirjalili and Lewis, WOA provide better performance as compared with PSO, GSA, GA, etc. as it has been tested with 29 mathematical benchmark and 6 structural design problem [18]. This method also has been successfully applied in many field such as machinery diagnosis [19,20], brain tumour [21], and breast cancer and tumour [22,23].

The rest of this paper is organized as follow. Section 2 presented the algorithm theory of ELM and WOA-ELM. Section 3 presented the WOA-ELM diagnosis approach and feature extraction using EMD and CEEMD. Section 4 draw the conclusion of the study.

* Corresponding author: firdaus060609@gmail.com

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).

2 Introduction to ELM and WOA-ELM

2.1 Extreme learning machine (ELM)

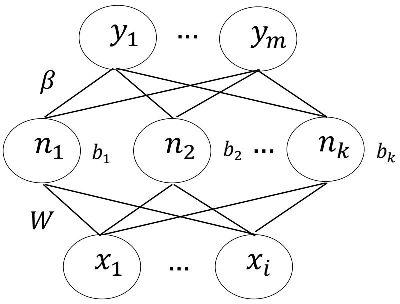

Figure 1. SLFN structural design.

The ELM was developed precisely only for handling SLFN structural design. Figure 1 shows the structure of SLFN where x is the input data array, W is the input weight, b is the hidden layer bias, β is the output weight and y is the output layer array. The annotation present as follows:

$$x = [x_1, \dots, x_i, 1] \qquad W = \begin{bmatrix} w_{11} & \cdots & w_{1k} \\ \vdots & \ddots & \vdots \\ w_{i1} & \cdots & w_{ik} \\ b_1 & \cdots & b_k \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_{11} & \cdots & \beta_{1s} \\ \vdots & \ddots & \vdots \\ \beta_{k1} & \cdots & \beta_{ks} \end{bmatrix} \qquad y = [y_1, \dots, y_s] \tag{1}$$

where i, k, and m are the input number, hidden neuron and output neuron. For computation ease, it has been observed that the weight and bias are placed together in the same matrix.

In ELM, input weight W is initialized randomly by sampling the values in the sense of uniform distribution between -1 and 1 and it is not change throughout the training phase. Analytical process of maintaining the universal approximation capability of SLFNs allow the construction of matrix β [9]. Then, hidden neurons feature map (H) is computed as follows:

$$h^j = [x_1^j, \dots, x_i^j, 1] \times \begin{bmatrix} w_{11} & \cdots & w_{1k} \\ \vdots & \ddots & \vdots \\ w_{i1} & \cdots & w_{ik} \\ b_1 & \cdots & b_k \end{bmatrix} \to H = \begin{bmatrix} f(h^1) \\ f(h^2) \\ \vdots \\ f(h^N) \end{bmatrix}_{N \times K} \tag{2}$$

where $f(.)$ is the activation function, $j = (1, \dots, N)$ and N is the number of samples in training dataset. Then, solve linear system using simple generalized inverse operation in order to compute β, as follow:

$$H\beta = Y \to \beta = H^{\dagger} Y \tag{3}$$

Where $H^{\dagger}$ is the Moore-Penrose generalized inverse of H. The Moore-Penrose solution is a least-square solution for a general linear system. Hence, it enable the ELM to achieve small training error, small norm of the weight and better generalization performance. Also, it does not stuck in local minima as the gradient descent-based learning methods [9]. The pseudocode of ELM algorithm presented as follow:

Step 1: *Function ELM_train* (X, Y, k, W, b).
Step 2: *Input: A training set* $\{X, Y\}$*, number of hidden neurons* k*, the input weight,* W*, and hidden layer bias* b.
Step 3: *If* W *and* b *are empty, then initialize* W *and* b *randomly between [-1, 1] and end.*
Step 4: *Compute feature map H from Equation (2).*
Step 5: *Compute* β *from Equation (3).*
Step 6: *Return:* W*,* H *and* β.

2.2 Parameter optimization using WOA-ELM

As mentioned, initialize the input weight, W and hidden layer bias, b will influence classification result. Optimization using meta-heuristic algorithm is one of the approach in order to determine the optimized value of W and b. In this study, a new and nature-inspired meta-heuristic method known as WOA is used with ELM to determine its parameter. Number of hidden neuron, n is also one of parameter for ELM. It do not has specific rule in determining the value but it also influence the classification result. Hence, there are three ELM parameters will be optimized in this study which are W, b, and n.

WOA was inspired by humpback whale hunting behaviour, also known as bubble-net hunting strategy [18]. This hunting strategy mainly consists of search of prey, encircling prey and bubble-net feeding manoeuvre. The complete description of each strategy is elaborated fully in reference [18].The pseudocode of WOA algorithm presented as follow:

```
Step 1: Initialize the whales population X_i (i = 1,2,...,n)
Step 2: Calculate the fitness of each search agent, X* is
the best search agent.
Step 3: while (t < maximum number of iterations)
    for each search agent
    Update a, A, C, l, and p
        if1 (p <0.5)
            if2 (|A| <1)
                Update position search agent
            else if2 (|A| ≥1)
                select random search agent
                update position search agent
            end if2
        else if1 (p ≥0.5)
            update position search agent
        end if1
    end for
    Check if any search agent goes beyond search space
    and amend it.
    Update X* if there is a better solution
    t = t + 1;
    end while
Step 4: return X* for best result.
```

The complete algorithm for parameter optimization using WOA-ELM shown in Figure 2. All three parameter has interval value set from [0,1] for W and b and [1,200] for n. The number of iteration is set to 100 with search agent number is set to 150. For ELM, "rbf" activation function is used. For objective function of WOA, the average classification accuracy of training and testing is used as described in Equation (4).

$$Fobj = \frac{Accuracy_{train} + Accuracy_{test}}{2} \tag{4}$$

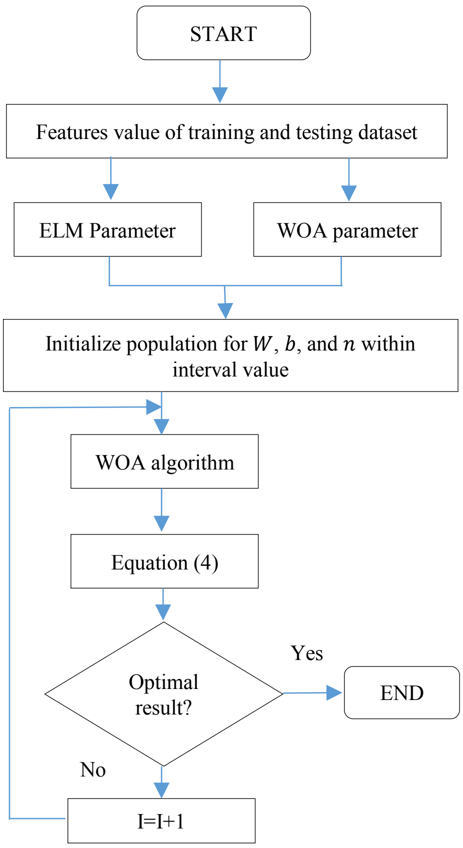

Figure 2: Proposed WOA-ELM method.

3 WOA-ELM for gearbox Diagnosis

3.1 Signal processing and feature extraction

As mentioned earlier, an online gearbox vibration signals downloaded from Acoustic and Vibration Database provided by Eric Bechhoefer are used [24]. The signals consists of healthy and faulty signal condition. The sampling rate for the signals is 97 656 Hz and it was recorded for 6 seconds. The signals was taken using accelerometer sensor in radial direction. Generally, gearbox vibration signals always subject to high noise contamination which make the extraction of features directly is not efficient [25]. Hence, empirical mode decomposition (EMD) and complementary ensemble empirical mode decomposition (CEEMD) have been used. These signal decomposition method at some extend has the capability to remove some noise from the vibration signals [26].

At first, all the vibration signals will be decomposed into sets of IMFs using EMD and CEEMD. Then, the statistical features listed in Table 1 are extracted from the combination of selected IMFs. The IMF was selected based on correlation criterion [27]. Then, the statistical features divided into training and testing dataset for classification. WOA is used to determine the optimized W, b, and n value for ELM algorithm. Full procedure of WOA-ELM for gearbox diagnosis shown in Figure 3.

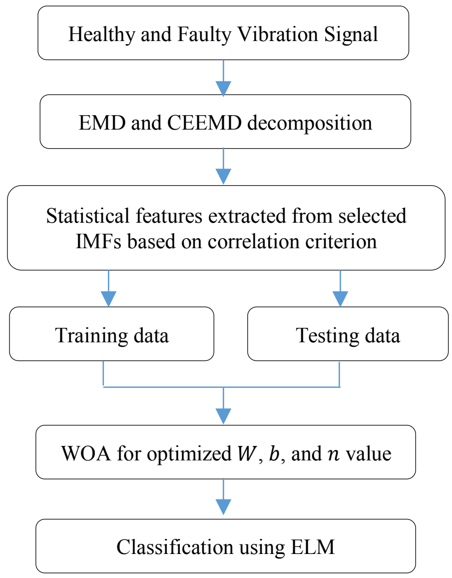

Figure 3: Gearbox diagnosis approach using WOA-ELM.

Table 1: Statistical features.

Statistical parameter	**Equation**
RMS	$\sqrt{\frac{\sum_{n=1}^{N} x(n)^2}{N}}$
Range	$\max(x) - \min(x)$
Skewness	$\frac{\frac{1}{N}\sum_{n=1}^{N}(x(n)-\bar{x})^3}{\left(\sqrt{\frac{1}{N}\sum_{n=1}^{N}(x(n)-\bar{x})^2}\right)^3}$
Kurtosis	$\frac{\frac{1}{N}\sum_{n=1}^{N}(x(n)-\bar{x})^4}{\left(\sqrt{\frac{1}{N}\sum_{n=1}^{N}(x(n)-\bar{x})^2}\right)^4}$
Crest factor	$\frac{max\lvert x(n)\rvert}{\sqrt{\frac{1}{N}\sum_{n=1}^{N} x(n)^2}}$
Shape factor	$\frac{\sqrt{\frac{1}{N}\sum_{n=1}^{N} x(n)^2}}{\frac{1}{N}\sum_{n=1}^{N}\lvert x(n)\rvert}$
Impulse factor	$\frac{max\lvert x(n)\rvert}{\frac{1}{N}\sum_{n=1}^{N}\lvert x(n)\rvert}$
Margin factor	$\frac{max\lvert x(n)\rvert}{\left(\frac{1}{N}\sum_{n=1}^{N}\sqrt{\lvert x(n)\rvert}\right)^2}$

The statistical features then divided into training and testing dataset. For ELM, all the features value need to be normalized within 0 to 1. Hence, the normalization of

 https://doi.org/10.1051/matecconf/201925502003

features are based on Equation (5). The distribution of these dataset summarized in Table 2.

$$S_i = \frac{s_i}{s_{i,max} \times C} \tag{5}$$

where S_i is a normalized features, s_i is statistical features, $s_{i,max}$ is the maximum statistical feature value and C is a constant for normalization which has been set to 1.3. Figure 4-11 shows the distribution of statistical features for EMD and CEEMD.

Table 2: Distribution of training and testing dataset for EMD and CEEMD features.

Dataset	Samples	Condition	Label
Training	45×8	Healthy	0
	45×8	Faulty	1
Testing	15×8	Healthy	0
	15×8	Faulty	1

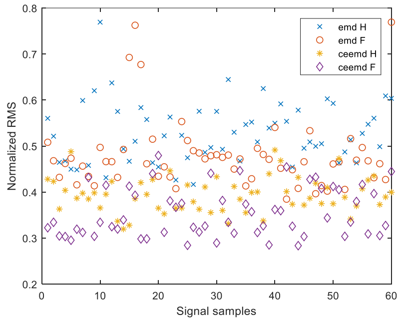

Figure 4: RMS features.

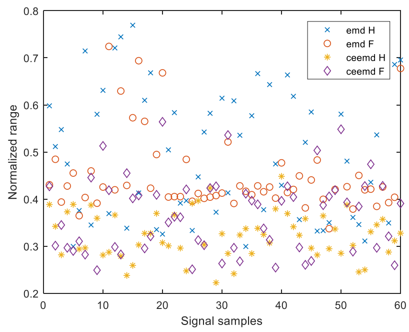

Figure 5: Range features.

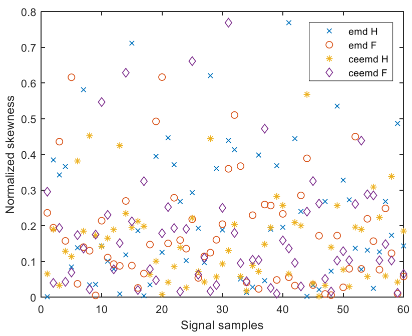

Figure 6: Skewness features.

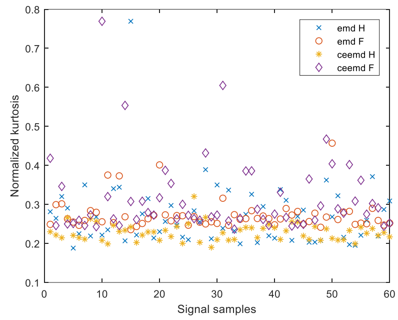

Figure 7: Kurtosis features.

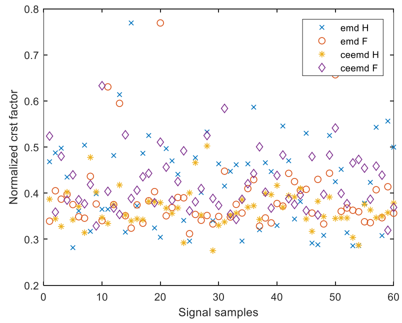

Figure 8: Crest factor features.

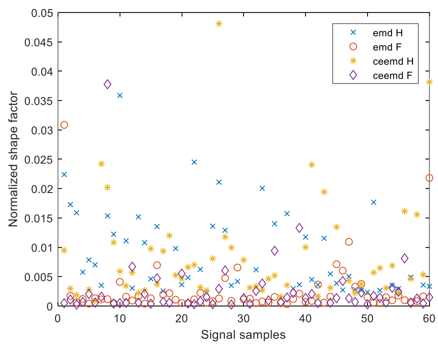

Figure 9: Shape factor features.

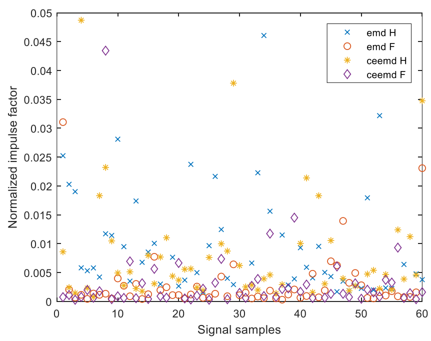

Figure 10: Impulse factor features.

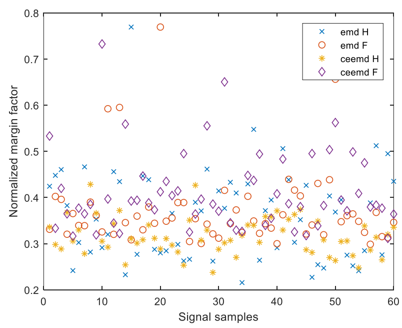

Figure 11: Margin factor features.

3.2 W, b, and n optimization using WOA-ELM

The value of W, b, and n will be optimized based on WOA-ELM. As mentioned above, the iteration is set to 100 and no of search agent is set to 150. The value of W, b, and n are set to its specific interval [0,1] for W and b and [1,200] for n. Figure 12 and 13 shows the optimization fitness convergence curve result for EMD and CEEMD.

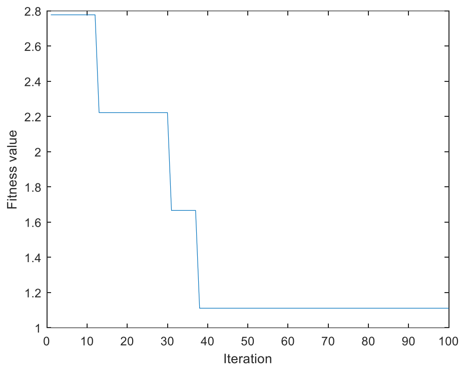

Figure 12: WOA-ELM result for EMD.

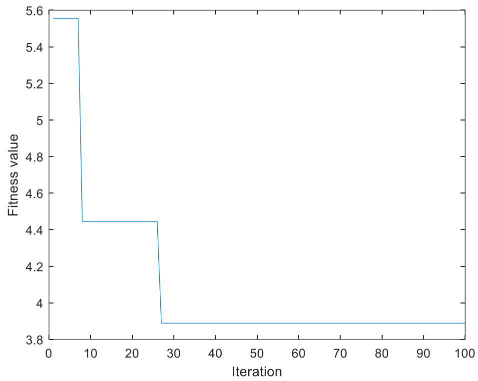

Figure 13: WOA-ELM result for CEEMD.

Based on these figure, the fitness value is converge early at iteration 38 for EMD and iteration 27 for CEEMD. The optimized value for W_{opt}, b_{opt}, and n_{opt} of EMD and CEEMD are [0.0848, 0.0974, 63] and [0.5919, 0.4544, 68]. Then, the value for W and b need to be expand into matrix. Hence, the optimized value will be acted as a middle value for the matrix where all the W value will be uniformly random within specific range of W_{opt}. It been described as Equation (6) and (7) below.

$$W_{opt,min} = \frac{W_{opt}}{3} \tag{6}$$

$$W_{opt,max} = W_{opt} \times 3 \tag{7}$$

The b value also expand into matrix by using the same concept, described in Equation (8) and (9):

$$b_{opt,min} = \frac{b_{opt}}{3} \tag{8}$$

$$b_{opt,max} = b_{opt} \times 3 \tag{9}$$

3.3 Classification based on WOA-ELM

By using the matrix W_{opt} and b_{opt} with n_{opt}, the classification is done. For running the ELM, it has been run for 30 times in order to avoid tweak problem and get a stable classification accuracy. A comparison between gearbox diagnosis result between WOA-ELM and conventional ELM have been studied. For conventional ELM, the W and b are set in random and default way and n is set to 8 which is equivalent to the number of statistical features used. Its shows that the proposed WOA-ELM method are able to improve the classification accuracy which means tuning an optimized value for W, b, and n are very important. Figure 14 shows the training accuracy, Figure 15 shows the testing accuracy and Figure 16 shows the average accuracy of training and testing. Based on the result, the proposed WOA-ELM is able to improve the accuracy performance by 6% to 8% as compared with conventional ELM.

Table 3: Classification result of WOA-ELM and conventional ELM.

Approach	Training (%)	Testing (%)	Overall (%)
EMD-ELM	83.41	81.44	**82.43**
EMD-WOA-ELM	97.78	76.67	**87.22**
CEEMD-ELM	85.33	88.33	**86.83**
CEEMD-WOA-ELM	97.78	86.67	**92.22**

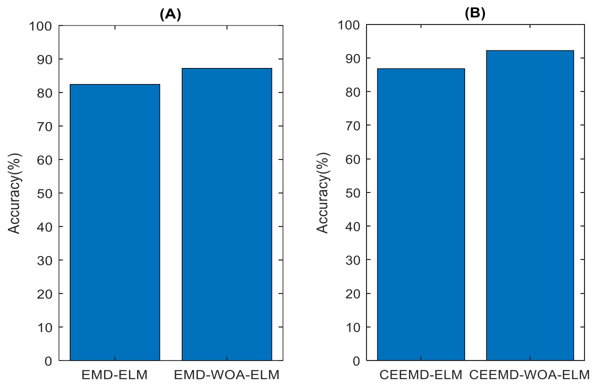

Figure 14: Overall classification accuracy, (A) using EMD features and (B) using CEEMD features.

4 Conclusions

This paper present the parameter optimization using WOA for determine the optimized input weight, bias and number of hidden neurons for ELM algorithm, based on MATLAB platform. The proposed method has been applied to the gearbox diagnosis study. An online gearbox vibration signals are used in this study. EMD and CEEMD method have been used for signal processing and features extraction procedure. A comparison has been presented between proposed WOA-ELM methods with conventional ELM method in this paper. The diagnosis accuracy shows that WOA-ELM provide better accuracy as compared with conventional ELM where the input weight and bias is set to random value and hidden neuron is set equivalent to the number of features extracted. Therefore, applying WOA-ELM for gearbox diagnosis can improve the failure efficiency. For future study, the proposed WOA-ELM will be applied in other rotating machinery applications such as bearing, shaft and belt to verify the superiority of this method in machinery diagnosis study.

Acknowledgement

The authors would like to extend their greatest gratitude to the Institute of Noise and Vibration UTM for funding the study under the Higher Institution Centre of Excellence (HICoE) Grant Scheme (R.K130000.7809.4J226). Additional funding for this research also comes from the UTM Research University Grant (Q.K130000.2543.11H36), and Fundamental Research Grant Scheme (R.K130000.7840.4F653) by The Ministry of Higher Education Malaysia. Besides that, the author also want to acknowledge Eric Bechhoefer for providing the gearbox dataset from Vibration and Acoustic Database.

References

[1] H. Ebrahimi Orimi, M. Esmaeili, A. Refahi Oskouei, S.A. Mirhadizadehd, P.W. Tse, Defect detection of helical gears based on time–frequency analysis and using multi-layer fusion network, Nondestruct. Test. Eval. 9759 (2016) 1–18. doi:10.1080/10589759.2016.1254211.

[2] J.M. Ha, B.D. Youn, H. Oh, B. Han, Y. Jung, J. Park, Autocorrelation-based time synchronous averaging for condition monitoring of planetary gearboxes in wind turbines, Mech. Syst. Signal Process. 70–71 (2016) 161–175. doi:10.1016/j.ymssp.2015.09.040.

[3] M. Nie, L. Wang, Review of condition monitoring and fault diagnosis technologies for wind turbine gearbox, Procedia CIRP. 11 (2013) 287–290. doi:10.1016/j.procir.2013.07.018.

[4] V. Sharma, A. Parey, A Review of Gear Fault Diagnosis Using Various Condition Indicators, Procedia Eng. 144 (2016) 253–263. doi:10.1016/j.proeng.2016.05.131.

[5] K.H. Hui, M.H. Lim, M.S. Leong, S.M. Al-Obaidi, Dempster-Shafer evidence theory for multi-bearing faults diagnosis, Eng. Appl. Artif. Intell. 57 (2017) 160–170. doi:10.1016/j.engappai.2016.10.017.

[6] M.F. Isham, M.S. Leong, M.H. Lim, Z.A. Ahmad, Variational Mode Decomposition for Rotating Machinery Condition Monitoring Using Vibration Signals, Trans. Nanjing Univ. Aero. Astro. 35 (2018) 38–50. doi:10.16356/j.1005-1120.2018.01.038.

[7] S.M. Ali, K.H. Hui, L.M. Hee, M.S. Leong, Automated valve fault detection based on acoustic emission parameters and support vector machine, Alexandria Eng. J. (n.d.). doi:http://doi.org/10.1016/j.aej.2016.12.010.

[8] L. Barbini, A.P. Ompusunggu, A.J. Hillis, J.L. du Bois, A. Bartic, Phase editing as a signal pre-processing step for automated bearing fault detection, Mech. Syst. Signal Process. 91 (2017) 407–421. doi:http://dx.doi.org/10.1016/j.ymssp.2016.12.004.

[9] G. Bin Huang, Q.Y. Zhu, C.K. Siew, Extreme learning machine: Theory and applications, Neurocomputing. 70 (2006) 489–501. doi:10.1016/j.neucom.2005.12.126.

[10] Y. Tian, J. Ma, C. Lu, Z. Wang, Rolling bearing fault diagnosis under variable conditions using LMD-SVD and extreme learning machine, Mech. Mach. Theory. 90 (2015) 175–186. doi:https://doi.org/10.1016/j.mechmachtheory.2015.03.014.

[11] Q. Tong, J. Cao, B. Han, X. Zhang, Z. Nie, J. Wang, Y. Lin, W. Zhang, A Fault Diagnosis Approach for Rolling Element Bearings Based on RSGWPT-LCD Bilayer Screening and Extreme Learning Machine, IEEE Access. 5 (2017) 5515–5530. doi:10.1109/ACCESS.2017.2675940.

[12] M.S.R.M. Saufi, Z.A. Ahmad, M.H. Lim, M.S. Leong, A review on signal processing techniques for bearing diagnostics, Int. J. Mech. Eng. Technol. 8 (2017) 327–337.

[13] D. Wang, S. Wei, H. Luo, C. Yue, O. Grunder, A novel hybrid model for air quality index forecasting based on two-phase decomposition technique and modified extreme learning machine, Sci. Total Environ. 580 (2017) 719–733. doi:10.1016/j.scitotenv.2016.12.018.

[14] Y. Chen, M. Kloft, Y. Yang, C. Li, L. Li, Mixed kernel based extreme learning machine for electric load forecasting, Neurocomputing. 312 (2018) 90–106. doi:10.1016/j.neucom.2018.05.068.

[15] T. Benkedjouh, S. Rechak, Intelligent prognostics based on empirical mode decomposition and extreme learning machine, Model. Identif. Control (ICMIC), 2016 8th Int. Conf. (2016) 943–947.

[16] Y. Wang, F. Cao, Y. Yuan, A study on effectiveness of extreme learning machine, Neurocomputing. 74 (2011) 2483–2490. doi:https://doi.org/10.1016/j.neucom.2010.11.030.

[17] D. Wang, P. Wang, Y. Ji, An oscillation bound of the generalization performance of extreme learning machine and corresponding analysis, Neurocomputing. 151 (2015) 883–890. doi:https://doi.org/10.1016/j.neucom.2014.10.006.

[18] S. Mirjalili, A. Lewis, The Whale Optimization Algorithm, Adv. Eng. Softw. 95 (2016) 51–67. doi:10.1016/j.advengsoft.2016.01.008.

[19] J. Xu, Y. Wang, Z. Ji, Fault Diagnosis Method of Rolling Bearing Based on WKELM Optimized by Whale Optimization Algorithm, Xitong Fangzhen Xuebao / J. Syst. Simul. 29 (2017) 2189–2197.

doi:10.16182/j.issn1004731x.joss.201709042.
[20] X. Zhang, Z. Liu, Q. Miao, L. Wang, Bearing fault diagnosis using a whale optimization algorithm-optimized orthogonal matching pursuit with a combined time–frequency atom dictionary, Mech. Syst. Signal Process. 107 (2018) 29–42. doi:10.1016/j.ymssp.2018.01.027.
[21] Virupakshappa, B. Amarapur, Computer-aided diagnosis applied to MRI images of brain tumor using cognition based modified level set and optimized ANN classifier, Multimed. Tools Appl. (2018) 1–29. doi:10.1007/s11042-018-6176-1.
[22] G.I. Sayed, A. Darwish, A.E. Hassanien, J.-S. Pan, Breast cancer diagnosis approach based on meta-heuristic optimization algorithm inspired by the bubble-net hunting strategy of whales, Adv. Intell. Syst. Comput. 536 (2017) 306–313. doi:10.1007/978-3-319-48490-7_36.
[23] M. Abdel-Nasser, A. Saleh, A. Moreno, N.S. Tabalvandani, D. Puig, Feature learning for breast tumour classification using bio-inspired optimization algorithms, Front. Artif. Intell. Appl. 300 (2017) 106–115. doi:10.3233/978-1-61499-806-8-106.
[24] E. Bechhoefer, High Speed Gear Dataset, Acoust. Vib. Database. (2014). http://data-acoustics.com/measurements/gear-faults/gear-1/.
[25] Y. Lei, D. Han, J. Lin, Z. He, Planetary gearbox fault diagnosis using an adaptive stochastic resonance method, Mech. Syst. Signal Process. 38 (2013) 113–124. doi:10.1016/j.ymssp.2012.06.021.
[26] R. Shao, W. Hu, J. Li, Multi-fault feature extraction and diagnosis of gear transmission system using time-frequency analysis and wavelet threshold de-noising based on EMD, Shock Vib. 20 (2013) 763–780. doi:10.3233/SAV-130783.
[27] M. Firdaus Isham, M. Salman Leong, L.M. Hee, Z.A.B. Ahmad, Empirical mode decomposition: A review on mode selection method for rotating machinery diagnosis, Int. J. Mech. Eng. Technol. 8 (2017).