Progress in Engineering Application and Technology Vol. 2 No. 1 (2021) 642–652
© Universiti Tun Hussein Onn Malaysia Publisher's Office

PEAT

Homepage: http://penerbit.uthm.edu.my/periodicals/index.php/peat
e-ISSN : 2773-5303

Development of Detection and Prediction System for Induction Motor Faults using Linear Regression

Nurul Nadhirah Mohd Sani[1], Faridah Hanim Mohd Noh[1]*

[1]Department of Electrical Engineering Technology, Faculty of Engineering Technology,
Universiti Tun Hussein Onn Malaysia, 84600 Pagoh, Johor, MALAYSIA

*Corresponding Author Designation

DOI: https://doi.org/10.30880/peat.2021.02.01.062
Received 13 January 2021; Accepted 01 March 2021; Available online 25 June 2021

Abstract: Induction motors (IM) are one of the commonly used electrical machines in the industry because of its greatness in the performance. They face various stresses during operating conditions that lead of occurring unexpected faults. In order to avoid the risk of unexpected failures and to improve the performance of the induction motor, condition monitoring becomes necessary. Hence, condition monitoring of the machines has become a more important strategy in structural health monitoring (SHM) research. Therefore, the aim of this study is to perform fault analysis of the induction motor and to study the failure identification techniques. The analysis for fault identification is categorized into two significant main components, which are feature extraction and prediction of the error. The first is used to extract the data from the signal and used to make prediction for the error in the IM. This paper uses a combination between Empirical Wavelet Transform (EWT) and Linear Regression (LR), in order to develop effective strategic method for detection and prediction using programming code in MATLAB software. The wavelet filter bank is created by EWT in order to extract the amplitude modulated-frequency modulated component of signal. The time domain features (TDFs) are then added to the reconstructed signal to extract the fault features. Then, it be the data for LR to developed and categorize the error and coefficient of induction motor faults. Table 1 shows the coefficient of determination, R^2 for bearing fault. As observed from Table 1, R^2 for the spectral kurtosis is 0.9955, for spectral crest factor is 0.9967 and for spectral entropy is 0.9986. The coefficient of determination is quite similar for the various features but in details, the coefficient of entropy is higher which is almost closed to ideal value of 1. It indicated an excellent coefficient if the value of R^2 is close or same value with 1. The experimental results indicate that, under various condition, the techniques can reliably extract and diagnose the IM fault. In addition, the efficiency of the EWT and LR indicates the superiority of the techniques suggested.

Keywords: Induction Motor (IM), Feature Extraction, Empirical Wavelet Transform (EWT), Linear Regression (LR), Time Domain Features (TDFs)

*Corresponding author: hanim@uthm.edu.my

1. Introduction

Three phase induction motor is the most important electrical equipment and widely used in the industry. It is more popular in industry since it can be used in many applications due to the uncomplicated structure, low maintenance cost and easy to control, as well as high reliability [1]. The importance of induction motor is proved by the fact that all factories require machinery in their manufacturing processes.

Despite the high reliability of the induction motor, there is still a possibility that the induction motor may have failures. The occurrence of motor faults might be due to mechanical and electrical stresses. There are numerous faults that occurred in the induction motor such as bearing faults, air-gap eccentricity, broken rotor bar, and the stator short winding. Some researchers have made an outcome that 30.00 -40.00 % of all reported failures in the stator or armature triggered by shortening of the stator winding phase and 5.00 -10.00 % failure in the split rotor bar and/or end ring fault [2]. The occurrence of these faults in the induction motor will degrade its performance and can lead to the destruction of the motor.

In finding an efficient and reliable fault diagnostic technique of an induction motor, numerous method can be proposed, where some of the method are the model based, signal based and smart method [3]. We can refer to Supervised Learning algorithm in which smart method that called as Regression Learning. Due to its simplicity and widespread use in the past for related topics, linear regression has been chosen as the method used in this research to predict the validity and precision of extracted features from faulty and normal signal of an induction motor. Linear regression is a way to predict a dependent variable’s relationship to one or more independent variables [4]. The linear regression helps to map numeric inputs to numeric outputs and fitting well a line into the data points and predicted the continuous quantity. The root means square error (RMSE) and the coefficient of determination are calculated in order to determine the amount of error within the model [5].

In this study, the main target is to apply the EWT and linear regression technique in detecting and predicting the induction motor fault. The effectiveness of extracted features is analyzed from Empirical Wavelet Transform (EWT) based on Root Mean Square Error (RMSE) and the coefficient of determination (R^2) from Linear Regression algorithm. The proposed algorithm is developed and analyze via MATLAB/Simulink environment.

This project focuses on collecting the raw data of the bearing faults of a three-phase induction motor that include of vibration signal. Then, the data will be decomposed using EWT algorithm, and the signal be the inputs for the feature extraction in time domain which was developed using MATLAB/Simulink software. This research target on the detection and evaluated root mean squared error (RMSE) with the coefficient of determination to quantify the error in the induction motor.

2. Literature Review

2.1 Induction Motor Fault

While induction motors are dependable electric machines, they are easily affected to many types of faults, include inter-turn short circuits in stator windings, broken rotor bars, bearing failures and rotor eccentricities. These failures can give effect on induction motors include unbalanced stator voltages and currents, torque fluctuations, reduced performance, overheating, excessive vibration and reduced torque [6]. Reportedly, nearly 40.00 % of total machine failures are due to bearing failure and the rates of failure for stator are reported as 47.00 %, for rotor failures as 5.00 %, for carriage defects as 32.00 % and for other defects as 16.00 %.

2.2 Vibration Monitoring Technique

There are several condition control methods that have been utilized but the oldest tracking method is vibration control. Since it is a precise and reliable tool for a machine health control measure, it is commonly used to identify bearing defects or mechanical imbalances. Vibration signals have non-deterministic and non-stationary properties that may render this vibration signal more effective in determining the motor condition of the induction and assessing the overall protection of a rotor device. Hence the induction motor vibration signal is used to distinguish between defective data and regular data. Researchers then used this vibration for fault detection and vibration feature to diagnose different faults in a very efficient way, in specific bearing faults, rotor eccentricities, gear faults and unbalanced rotors [6].

2.3 Time Domain Feature Extraction Analysis

The time-domain feature is based on waveform time-signal statistical aspects. The most frequently used features in time domain analysis include Root Mean Square (RMS), Crest Factor, Skewness, Kurtosis, Variance, and Standard Deviation. RMS is effective in time-domain applications but failing components can't be evaluated and the machinery fault not well understood. However, it has the ability to determine any fault with imbalances, particularly in rotating equipment [7]. Meanwhile, the crest factor (CF) will give the percentage of the peak of an input signal to the RMS stage and will be useful for detecting changes in the frequency of the impulses. Instead, in comparison to a normal distribution, kurtosis would provide the check for whether the data is elevated or smooth. However, there are still disadvantages which is lack of noticeable signs of faults especially while being a fault at an early stage and which is one of the limitations of the time domain features extraction technique. The technique may be successful when features of short duration are removed from the signal.

2.4 Empirical Wavelet Transform

The Empirical Wavelet Transform (EWT) is introduced to improve the accurate diagnosis rate. EWT is a technique of complete adaptive sub-band decomposition based on the maximum of given signal f(t) under the framework of wavelet transform theory. According to Gilles, he suggested this strategy to acquire adaptive wavelets capable of detecting and removing signal components from AM-FM [8]. The key idea is to use a fitting wavelet filter bank to extract the vibration signal modes. The Fourier frequency spectrum is normalized and divided into the number of intervals in the EWT process, then the closely supported orthonormal wavelet base is constructed directly at each interval [9]. The segmentation operation on the Fourier spectrum is then the critical step in making the wavelet adaptable to the signal.

2.5 Linear Regression

Nowadays, the regression learning has been applied in numerous areas. It is also widely used on forecasting and finding out cause and effect relationship between variables. According to Stanton by Sir Francis Galton, the regression theory was first popularized in the 19th century [9]. The analysis of regression system mostly contrasts based on the number of independent variables and the variety of relation between the independent and dependent variables. The utilization of the Linear Regression method is very easy to use and make it easily implement the regression method when the relationship of two variables had been identify. Then, the best fit line, which is used in prediction can be developed after performing linear regression. The output in the scatter graph of the linear regression that will show the best fit line. The simulation results will represent the root mean square values (RMSE) and the coefficient of determination in order to determine the error in the induction motor.

3. Methodology

The implementation of the feature extraction and prediction algorithm are based on the coding that created in MATLAB environment. Basically, this project consists of two main algorithms, which are features extraction algorithm and features prediction algorithm. Detail explanation on the stated algorithm will be explained in subtopic 3.1 to 3.4.

3.1 Flowchart of the overall systems

The process for the diagnosis techniques begins with the flowchart which is the step-by-step method that necessary for resolving the problem. By using this diagnosis technique, it was able to show the process flow in order to determine the error and solved the problem. The flowchart for the overall motor fault diagnosis techniques is shown below:

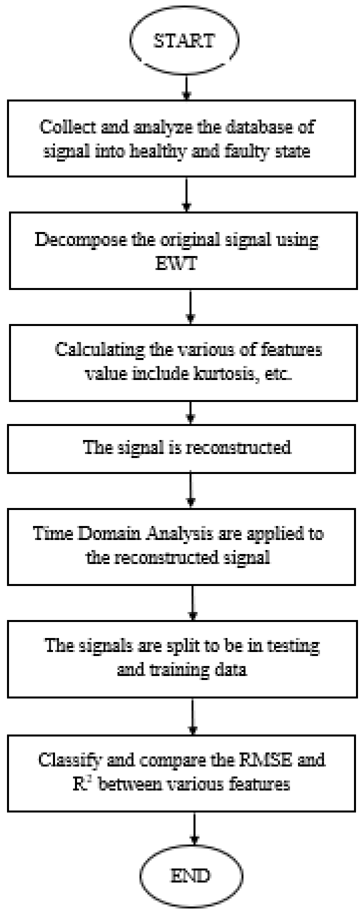

Figure 1: Flowchart of motor fault diagnosis techniques

The simulation was implemented by using MATLAB/Simulink software and the techniques above can be summarized as follows:

Step 1: Collecting the real signal database of vibration and current signal of the induction motor faults. The signal will be analyzed on two different states (healthy and faulty) in order to improve the input data.

Step 2: The original signal is decomposing into some band-pass frequencies using Empirical Wavelet Transform (EWT). After that, by calculating the kurtosis value, spectral entropy and crest factor, the signal is reconstructed from the significant modes.

Step 3: The time domain analysis is used to extract the health status of the faults in the induction motor for the reconstructed signal. Next, the features will split to the data testing with data training and it is selected to be the inputs for this system.

Step 4: After the selected features be the input, the linear regression functioned to determine the root means square error and the coefficient of determination based on the plotted linear regression graph under different features.

3.2 Features extraction using EWT

EWT is a new adaptive wavelet transform capable of decomposing a time series signal x(t) into adaptive time-frequency sub-bands according to its contained frequency information. This provide an adaptive wavelet with respect to the analyzed signal; the segmentation of the signal is an important step in the EWT. It can be performed either manually or by means of the Fourier spectrum. Figure 2 shows the basic construction for the EWT.

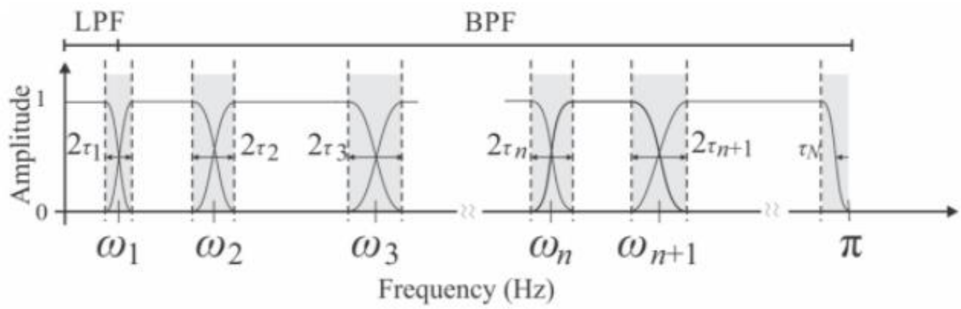

Figure 2: EWT basic construction [10]

The empirical scaling function to approximate the low-pass wavelet filter coefficients according to the following equation below and it is defined in accordance with the idea used in deriving the Meyer wavelet [10].

$$\emptyset_n(\omega) = \begin{cases} 1 \; if \; |\omega| \le \omega_n - \tau_n \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_n}(|\omega| - \omega_n + \tau_n)\right)\right] \; if \; \omega_n - \tau_n \le |\omega| \le \omega_n + \tau_n \\ 0 \; otherwise \end{cases} \qquad Eq. 1$$

$$\Psi_n(\omega) = \begin{cases} 1 \; if \; \omega_n + \tau_n \le \omega_{n+1} - \tau_{n+1} \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_{n+1}}(|\omega| - \omega_{n+1} + \tau_{n+1})\right) if \; \omega_{n+1} - \tau_{n+1} \le |\omega| \le \omega_{n+1} + \tau_{n+1}\right] \\ \sin\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_n}(|\omega| - \omega_n + \tau_n)\right) if \; \omega_n - \tau_n \le |\omega| \le \omega_n + \tau_n\right] \\ 0 \; otherwise \end{cases} \qquad Eq. 2$$

$$w_f^e(n,t) = F^{-1}\big(x(\omega)\Psi_n(\omega)\big) \qquad Eq. 3$$

$$w_f^e(0,t) = F^{-1}\big(x(\omega)\emptyset_n(\omega)\big) \qquad Eq. 4$$

For this purpose, each signal is decomposed into the 9 modes with the help of EWT like as shown in Figure 3 below. The selected modes are used for calculating the kurtosis, crest factor and entropy values. Kurtosis are calculated by the fourth modes value from the EWT, while crest factor from sixth modes and entropy from fifth modes. Hence, the features indicator is used to reconstruct the signal and the envelope spectrum is applied for diagnosing the fault.

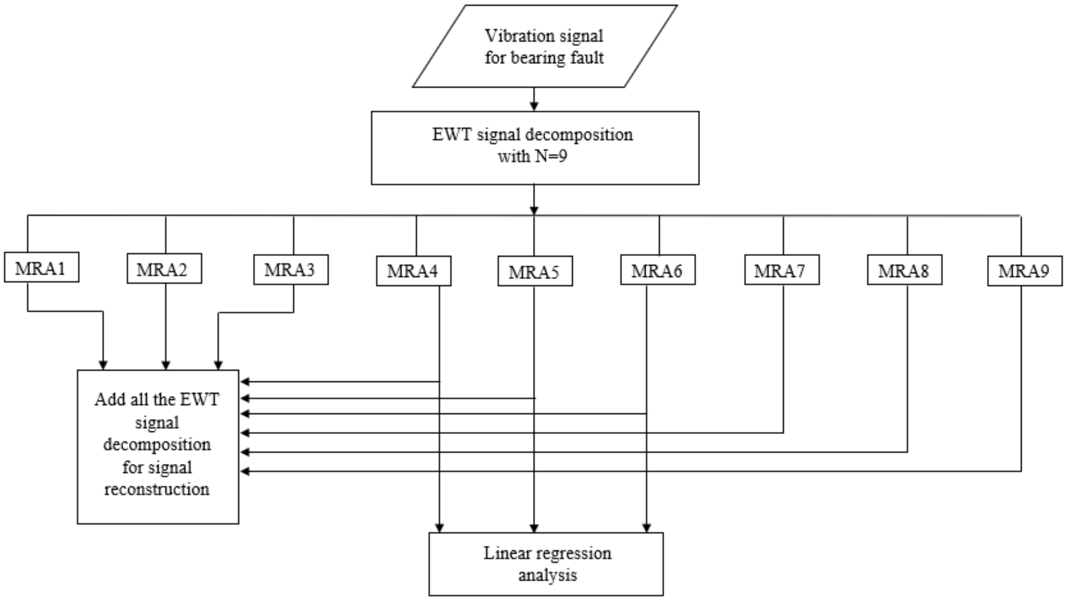

Figure 3: Flowchart of EWT technique

3.3 Feature prediction using Linear Regression

According to Stanton by Sir Francis Galton, the regression theory was first popularized in the 19th century [14]. The analysis of regression system mostly contrasts based on the number of independent variables and the variety of relation between the independent and dependent variables. The utilization of the Linear Regression method is very easy to use and make it easily implement the regression method when relationship of two variables had been identify. With the using feature that extracted by EWT and TD analysis, the output in the scatter graph of the linear regression will show the fit line between measured values and the target values. Then, the best fit line, which is used in prediction will represent the root mean square error values (RMSE) and the coefficient of determination in order to determine the error in the induction motor.

3.4 Numerical equations for extracted features

To extract the data of the signal from various states, the analytical features are calculated using the equations below.

$$\text{Entropy} = -\sum p \log(p) \qquad \textit{Eq. 5}$$

$$\text{Kurtosis} = \frac{E(x-\mu)^4}{\mu^4} \qquad \textit{Eq. 6}$$

$$\text{Crest factor} = \frac{max|x(n)|}{RMS} \qquad \textit{Eq. 7}$$

4. Results and Discussion

The simulation is purposely to find the performances of the fault detection system with the comparing RMSE values and R^2 values for different features. The simulation results are based on the working process for the fault diagnosis system according to the programming code in the MATLAB software.

4.1 Data description

Figure 4 and 5 below show the raw signal collected from induction motor at healthy and faulty data of bearing signal that was processed. Both graph signals below show in form of acceleration versus time.

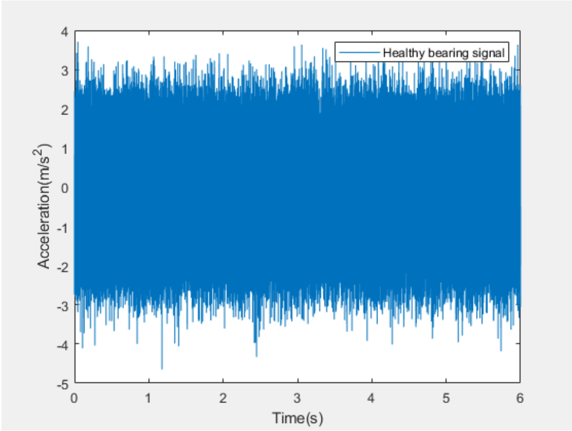

Figure 4: The induction motor signal for healthy bearing condition

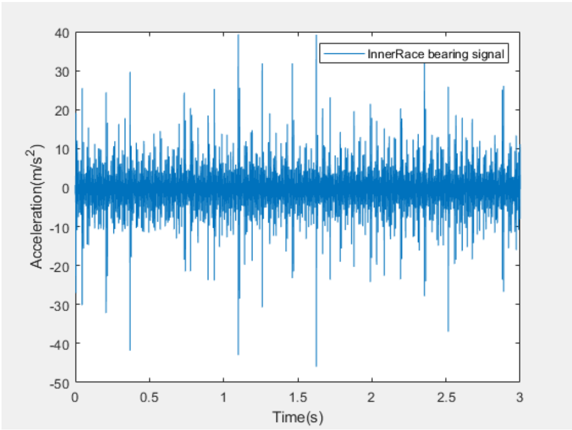

Figure 5: The induction motor signal for faulty bearing condition

4.2 Condition monitoring

The Empirical Wavelet Transform are used in this paper to decompose the original signal into several band-pass frequencies (modes). As can be seen from Figure 6, it shows the same set of EWT-processed raw signals with the number of component modes is 9 and appear completely distinguishable from each other.

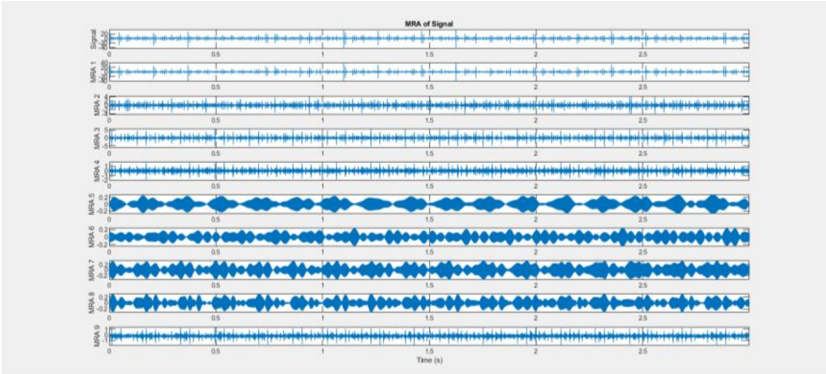

Figure 6: The extracted modes of the signal using EWT

In order to validate the effectiveness of the EWT system for simulated signal processing, the decomposed signals are reconstructed by calculating the kurtosis value, crest factor and entropy values, respectively. The reconstructed signal is shown in Figure 7.

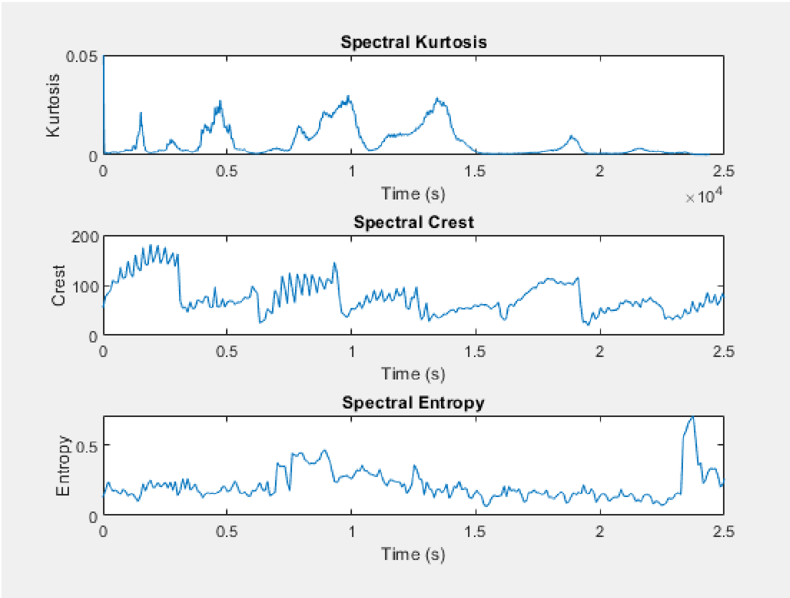

Figure 7: Reconstructed signal for bearing faults

As can be seen in Figure 7 above, it shows the clear extended excursion and to take a closer look based on the spectral graph for kurtosis, crest factor and entropy. Based on the graph, it is observed that the amplitude of the signal has its own significantly impulsiveness in different features and it allowing the fault signature easily captured by spectral diagram.

4.3 Prediction technique

As regression analysis is selected as the technique to construct the defect prediction, the fault bearing data collected are functional in the decomposed and reconstructed signal. Then, the selected features values from the Kurtosis, Crest Factor and Entropy are plotted in the scatter graph to show the best fitted line. However, Figure 8 to 10 shows that linear regression graph for Kurtosis, Crest Factor and Entropy in form of testing data vs expected values.

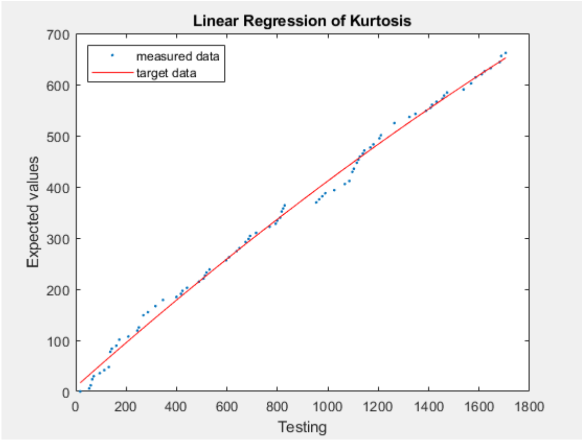

Figure 8: Actual and fitted Kurtosis

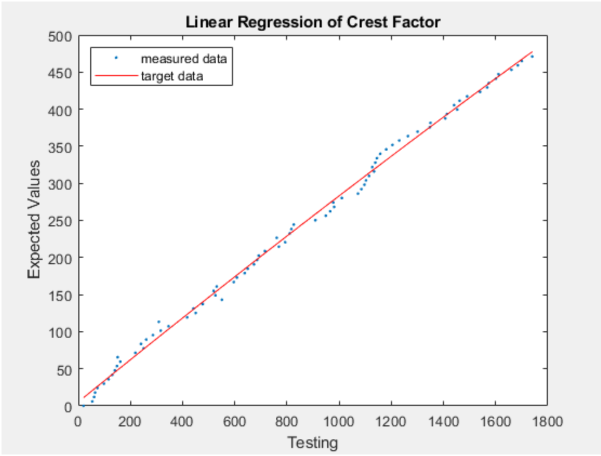

Figure 9: Actual and fitted Crest Factor

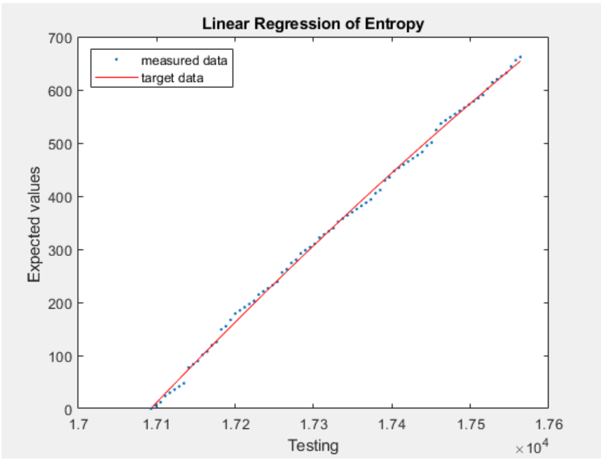

Figure 10: Actual and fitted Entropy

As can be seen in Figure 8 to 10 above, it seems that there is linear regression graph trend between the measured data and the target data. Based on all three results as represented from Figure 8 to 10, all the fitted line looked similar but the linear regression graph for entropy are show that the measured data more accurate with the target data compared to other features. This is because linear regression tries to find a straight line that best fits the data. However, it proved by the regression graph is not in a straight line and the scattered points of the plot are not closely with the line, it means that there is much error or that the error is high.

Table 1: RMSE and R^2 values in linear regression

	Spectral Kurtosis	Spectral Crest Factor	Spectral Entropy
RMSE	12.8630	7.9574	7.2047
R^2	0.9955	0.9967	0.9986

The performance of the detection and prediction is measured using the comparison between Root Mean Square Error (RMSE) and the coefficient of determination (R^2) in the Table 1. Table 1 shows the overall performance of RMSE values and R^2 values for bearing fault. While, it also shows the coefficient of determination, R^2 for bearing fault in the various features. As observed from Table 1, it is show that the RMSE were recorded and the spectral kurtosis is 12.8630, 7.9574 is for spectral crest factor and 7.2047 is for spectral entropy. It is show that RMSE for entropy is much better for the system. The coefficient of determination is quite similar for the various features but in details, the coefficient of entropy is higher which is almost closed to ideal value of 1. It indicated an excellent coefficient if the value of R^2 is close or same value with 1.

5. Conclusion

An effective methodology was presented to diagnose the bearing faults in a three-phase induction motor based on EWT and linear regression. This paper proposes are comparative study of different features extracted for fault prediction in the induction motor. Fault detection techniques consist of decomposing signal in EWT and reconstructed by calculating the kurtosis, crest factor and entropy. While, fault prediction using statistical and machine learning methods such as linear regression was carried out by programming code in MATLAB environment. Table 1 shows that the RMSE and R^2 can be obtained better in fault prediction. It can be observed that the spectral entropy are the most effective features in determining the RMSE and R^2. It can be concluded from the linear regression analysis that the techniques appear to be more useful in predicting faults. However, it also proved by the regression graph that show the best fit in a straight line and the scattered points of the plot.

Acknowledgement

The authors like to thank the Faculty of Engineering Technology, Universiti Tun Hussein Onn Malaysia for its support.

References

[1] H. F. Azgomi and J. Poshtan, "Induction motor stator fault detection via fuzzy logic," *2013 21st Iran. Conf. Electr. Eng. ICEE 2013*, no. May 2013, 2013, doi: 10.1109/IranianCEE.2013.6599711.

[2] G. Jose and V. Jose, "Fault diagnosis in induction motor using soft computing techniques," *ICACCS 2013 - Proc. 2013 Int. Conf. Adv. Comput. Commun. Syst. Bringing to Table, Futur. Technol. from Around Globe*, 2014, doi: 10.1109/ICACCS.2013.6938693.

[3] S. Nandi, H. A. Toliyat, and X. Li, "Condition monitoring and fault diagnosis of electrical motors - A review," *IEEE Trans. Energy Convers.*, vol. 20, no. 4, pp. 719–729, 2005, doi: 10.1109/TEC.2005.847955.

[4] J. Hallman, "A comparative study on Linear Regression and Neural Networks for estimating order quantities of powder blends," 2019.

[5] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature," *Geosci. Model Dev.*, vol. 7, no. 3, pp.

1247–1250, 2014, doi: 10.5194/gmd-7-1247-2014.

[6] S. A. Da Silva Vicente, R. Y. Fujimoto, and L. R. Padovese, "Rolling bearing fault diagnostic system using fuzzy logic," *IEEE Int. Conf. Fuzzy Syst.*, vol. 2, pp. 816–819, 2001, doi: 10.1109/fuzz.2001.1009080.

[7] S. Narwade, P. Kulkarni, and C. Y. Patil, "Fault detection of Induction Motor using Envelope Analysis," *Int. J. Adv. Comput. Res. (ISSN*, vol. 2, no. 7, pp. 258–262, 2013.

[8] W. Deng, S. Zhang, H. Zhao, and X. Yang, "A novel fault diagnosis method based on integrating empirical wavelet transform and fuzzy entropy for motor bearing," *IEEE Access*, vol. 6, pp. 35042–35056, 2018, doi: 10.1109/ACCESS.2018.2834540.

[9] N. Huang *et al.*, "Short-circuit fault detection and classification using empirical wavelet transform and local energy for electric transmission line," *Sensors (Switzerland)*, vol. 17, no. 9, pp. 1–24, 2017, doi: 10.3390/s17092133.

[10] J. Gilles, "Empirical Wavelet Transform," vol. 61, no. 16, pp. 3999–4010, 2013.

[11] X. Zhang and Z. Teng, "A subspace type incremental two-dimensional principal component analysis algorithm," *J. Algorithms Comput. Technol.*, vol. 14, 2020, doi: 10.1177/1748302620973531.

[12] S. Gambino, A. Laudani, and S. Mangiagli, "Seismicity pattern changes before the M = 4.8 aeolian archipelago (Italy) earthquake of august 16, 2010," *Sci. World J.*, vol. 2014, no. Figure 1, pp. 1–9, 2014, doi: 10.1155/2014/531212.

[13] W. James, "Possible Application to the Great Chilean Earthquake of May 22 , 1960," vol. 87, no. 2, pp. 2851–2859, 1982.

[14] J. M. Stanton and J. M. Stanton, "Galton , Pearson , and the Peas : A Brief History of Linear Regression for Statistics Instructors Galton , Pearson , and the Peas : A Brief History of Linear Regression for Statistics Instructors," vol. 1898, 2017, doi: 10.1080/10691898.2001.11910537.

[15] S. Blanco, A. Figliola, R. Q. Quiroga, O. A. Rosso, and E. Serrano, "Time-frequency analysis of electroencephalogram series. III. Wavelet packets and information cost function," *Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top.*, vol. 57, no. 1, pp. 932–940, 1998, doi: 10.1103/PhysRevE.57.932.

[16] S. Chen and Y. Chen, "Hardware Design and Implementation of a Wavelet De-Noising," no. Ic, pp. 26396–26414, 2015, doi: 10.3390/s151026396.

[17] P. C. Machine, "Simultaneous-Fault Diagnosis of Gearboxes Using Probabilistic Committee Machine," 2016, doi: 10.3390/s16020185.