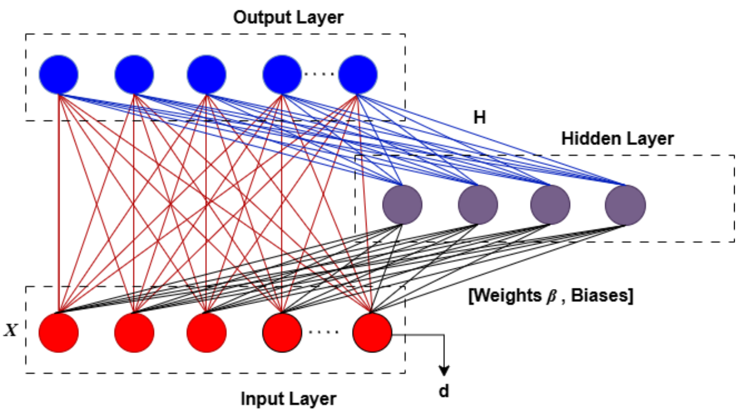
Type of contribution:
Editorial
▶ Research Paper
Case Study
Review Paper
Scientific Data
Report of Tech. Application

mesi
e-issn. 2798-5245
Mechanical Engineering
for Society and Industry
Vol. 5, No. 1 (2025) pp 190-205
https://doi.org/10.31603/mesi.12965

Remaining useful life prognosis of low-speed slew bearing using random vector functional link

Wahyu Caesarendra[1]*, Dimas Revindra Rahardja[2], Muhammad Abdillah[2], Seno Darmanto[3], Sri Utami Handayani[3], Wahyu Dwi Lestari[4], Grzegorz Krolczyk[1]

[1] Department of Mechanical and Mechatronics Engineering, Faculty of Engineering and Science, Curtin University Malaysia, 98009 Miri, Sarawak, **Malaysia**
[2] Department of Electrical Engineering, Pertamina University, Jakarta Selatan 12220, **Indonesia**
[3] Department of Industrial Technology, Vocational School, Diponegoro University, Semarang 50275, **Indonesia**
[4] Department of Mechanical Engineering, Faculty of Engineering, University of Pembangunan Nasional Veteran Jawa Timur, Surabaya 60294, **Indonesia**
✉ w.caesarendra@curtin.edu.my



Highlights:
- This article utilizes the Random Vector Functional Link (RVFL) method to predict the Remaining Useful Life (RUL) of slew bearing.
- This study presents a comparison of four activation functions i.e., SELU, ReLU, Sigmoid, and Sine.
- The optimal model was achieved using an 80:20 data split for training and testing, along with the SELU activation function.
- The developed model is capable of predicting the RUL of bearings with an accuracy of 94.24%.

Article info
Submitted:
2025-01-05
Revised:
2025-05-03
Accepted:
2025-05-24

Publisher
Universitas Muhammadiyah Magelang

Abstract
Bearings have a very important role in an industry. However, the cost of maintenance and replacement of bearings are very expensive, especially for slew-bearing which operates at a very low speed. If the low-speed slew bearing fault suddenly, it will shut down the entire rotating machine and also cause a financial issue due to the termination of production process in the certain industries. Therefore, monitoring of the low-speed slew bearing condition at all times is necessary to predict the bearing failure. There has been advance monitoring devices and systems related to the vibration condition monitoring for bearing and rotating machines, however, in certain cases those monitoring devices and systems are not sufficient due to a lack of decision support system. Machine learning (ML) is offered to complement and contribute in this case which aims to predict the Remaining Useful Life (RUL) severe damage occurred. In this paper, the Random Vector Functional Link (RVFL) is used to predict the RUL of the vibration bearing data collected from run-to-failure low speed slew bearing experiment. A few of activation functions such as Scaled Exponential Linear Unit (SELU), Rectified Linear Unit (ReLU), Sigmoid, and Sine were also studied to obtained the most appropriate prediction model. The selection of the best activation function for the prediction model is based on the evaluation matrix such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). According to the

prediction results, the best modeling results are obtained using a data ratio of 80:20 and the SELU activation function that produces the best average RMSE value. The RUL prediction of the bearing is 94.24%. In practice, this production accuracy is acceptable. However, further study to improve the accuracy is necessary by increase the bearing data as well as the sample of bearing under investigation. The RVFL method were also compared with Extreme Learning Machine (ELM) and Artificial Neural Network (ANN) with the result of RVFL is outperformed.

Keywords: Low speed slew bearing; Remaining useful Life; Random vector functional link

1. Introduction

For decades, technology has been developing rapidly, particularly in the industrial sector. This development has significantly impacted various industrial fields including the steel manufacturing industry. In the industrial sector and particularly steel manufacturing, rotating machinery plays an important role in supporting the production. In steel manufacturing, slew bearings operating at low rotational speeds are frequently utilized in the rotating machines. These bearings often encounter conditions that require them to perform under strenuous circumstances, resulting in high replacement costs and extended delivery times [1]. Consequently, continuous condition monitoring of slew bearings is essential [2], [3]. Typically, maintenance and replacement of unserviceable low-speed slew bearings are performed as preventive measures to avert unexpected failures. To optimize the use of low-speed slew bearings and prevent sudden breakdowns, early fault detection monitoring methods are required. One such method involves the prognosis of Remaining Useful Life (RUL) [4]. Prognosing the RUL of bearings is crucial for monitoring their condition, preventing undesirable downtime, and enhancing machine reliability [3]. The RUL of a bearing refers to the estimated remaining life before it experiences functional failure, which, in this design, is indicated by a degradation index value surpassing a defined threshold.

Predicting the RUL of bearings can be achieved using various machine learning (ML) methods [5], such as Linear Regression (LR) [6], Graph Neural Network (GNN), Random Forest (RF), Bayesian Regression (BR), and Random Vector Functional Link (RVFL). In addition, deep learning methods [7] such as Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) were recently applied in bearing fault diagnostics [8], [9].

In this study, the run to failure bearing data will be predicted using the RVFL method. The RVFL is a Single Layer Feedforward Neural Network (SLFNN) in which the weights and biases of the hidden neurons are randomly generated within an appropriate range and kept constant, while the output weights are computed through a simple closed-form solution [10]. Random-based neural networks benefit significantly from the direct connections between the input and output layers, as seen in RVFL networks.

This design offers the RVFL method to determine the RUL of bearings. The RVFL method offers several advantages over other ML methods, including rapid learning, simple design, high generalization capability, universal approximation ability, and efficient estimation accuracy [11], [12]. Based on these advantages, the RVFL method is offered in this design to predict and estimate the RUL of bearings.

2. Methods

2.1. Feature Extraction

The dataset utilized for the RVFL prediction consists of five time-series features: Root Mean Square (RMS), kurtosis, variance, histogram upper, and histogram lower. These features are extracted from the raw vibration data available and are used to monitor the daily condition of high-load slew bearings undergoing degradation. They serve as indicators of the bearing's state, as significant changes in these features correspond to increased vibration levels with associate to the bearing condition. When the damage becomes severe, the vibration signal surpasses the abnormal level. Therefore, it is essential to employ reliable and appropriate features for monitoring the condition of slew bearings. The selected time-series features for this study include is described in as follows:

2.1.1. Root Mean Square (RMS)

The RMS feature is the square root value of the average signal sum of squares as presented in Eq. (1).

$$RMS = \sqrt{\frac{1}{N}\sum\nolimits_{i=1}^{N} x_i^2} \tag{1}$$

where x_i are the individual vibration data values and N is the number of data values.

2.1.2. Variance

The Variance feature is used to measure the dispersion of a signal or dataset around its mean reference. The Variance feature is represented in Eq. (2).

$$\sigma^2 = \frac{1}{N}\sum\nolimits_{i=1}^{N} (x_i - m)^2 \tag{2}$$

where *m* is the mean of the vibration data.

2.1.3. Kurtosis

The Kurtosis feature is used to measure the flatness of the Probability Density Function (PDF) near the centre. The Kurtosis feature is calculated in Eq. (3).

$$Ku = \frac{1}{N}\sum\nolimits_{i=1}^{N} \left(\frac{x_i - m}{\sigma}\right)^4 \tag{3}$$

where σ is the standard deviation of the vibration data.

2.1.4. Histogram Upper

The histogram is a discrete probability density function. The Histogram Upper feature is described in Eq. (4).

$$\begin{gathered} h_i = \sum_{i=0}^{N} r_i(x_i) \\ h_U = \max(x_i) + \Delta/2 \end{gathered} \tag{4}$$

where h_i is the columns of the histogram for the time x_i with the range $0 \ll i < d$
where r_i is

$$r_i = \begin{cases} 1, & if \ \frac{i(\max(x_i) - \min(x_i))}{d} \le x < \frac{(i+1)(\max(x_i) - \min(x_i))}{d} \\ 0, & otherwise \end{cases} \tag{5}$$

and d be the number of divisions that group the ranges

2.1.5. Histogram Lower

The histogram is a discrete probability density function. The Histogram Lower feature is calculated in Eq. (6).

$$\begin{gathered} h_i = \sum_{j=0}^{n} r_i(x_i) \\ h_L = \max(x_i) - \Delta/2 \end{gathered} \tag{6}$$

2.2. Degradation Index

Degradation index is used as a part of prognostic method to predict maintenance and health monitoring of the bearing [13]. The degradation index value can be determined by summing all the features that will be used [14]. The degradation index is presented in Eq. (7):

$$Z = \frac{1}{r}\sum_{i=1}^{t} r_t \tag{7}$$

where *Z* is degradation index, *t* = time (day), and *r* = feature extraction

2.3. Random Vector Functional Link

Random Vector Functional Link (RVFL) is a part of ML that belongs to the supervised learning category. The RVFL is a single-layer feed-forward neural network where the weights and biases of the hidden neurons are randomly generated within an appropriate range and kept similar. In contrast, the output weights are calculated from a simple closed-form approach [10]. The direct link from the input layer to the output layer in RVFL networks makes the randomization-based neural networks advantageous [15]. The original features are reused and passed to the output layer through the direct connection. The direct link serves as a regularization for randomization. With the direct relationship between input and output, the model complexity of this RVFL network becomes lower and simpler when compared to other methods [10].

RVFL is a randomized version of the Single Layer Feedforward Neural Network or SLFNN, with three layers: input layer, hidden layer, and output layer. The three layers consist of neurons connected through weights [16]. The layers are interconnected as shown in RVFL structure and presented in **Figure 1**.

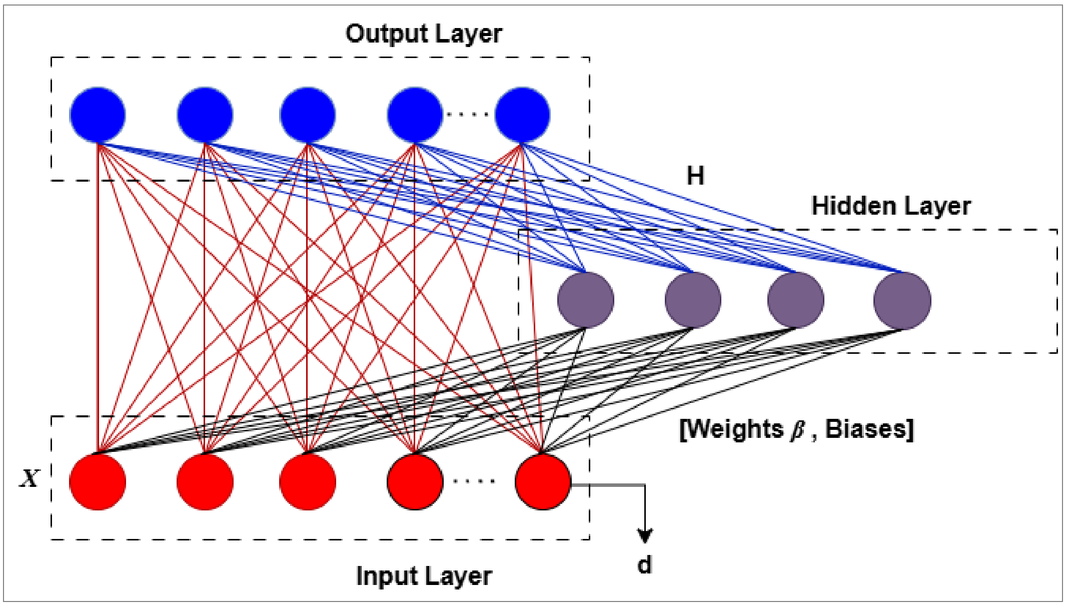


Figure 1. RVFL structure. Adapted from [14]

A detail structure of the RVFL method including the layers and the parameters is presented in **Figure 1**. According to **Figure 1**, the input layer to the output layer in RVFL consists of non-linearly transformed features H from the hidden layer and the original input features in *X*. If d is the input data features and *N* is the number of hidden nodes, then there is a total of d + *N* inputs for each output node. Since the matrix H are randomly generated parameters and fixed during the training phase, only the output weight *β* needs to be calculated. For the prediction purpose, the RVFL prediction model is calculated in Eq. (8) [16].

$$y_i = \sum_{j=1}^{L} \beta_j h_j(w_j^T, b_j, x) + \sum_{j=L+1}^{L+m} \beta_j x_j \tag{8}$$

where y_i is the prediction output, L is number of neurons in hidden layer, b_j is randomly selected bias w_j is randomly selected weights, β_j is weight between the output neuron and the *j* neuron, and m is number of input features, and h_j is the activation function. The values of the hidden node parameters b_j and w_j are set randomly, while the layer parameters output are calculated analytically to ensure the performance of the network. The mathematical formulation of the hidden layer and output layer weights can be written in Eq. (9).

$$H = \begin{bmatrix} h(w_1, b_1, x_1) & \cdots & h(w_L, b_l, x_1)\, x_1 \\ \vdots & \vdots & \vdots \\ h(w_1, b_1, x_N) & \cdots & h(w_L, b_l, x_N) x_N \end{bmatrix}_{N\,x\,M} \tag{9}$$

where *H* is matrix of hidden layer output.

The output weight matrix β and the target output prediction matrix Y is presented in Eqs. (10) and (11).

$$\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_M \end{bmatrix}_{M\ x\ 1} \tag{10}$$

$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}_{N\ x\ 1} \tag{11}$$

According to the mathematical formula derivation in Eqs. (9) – (11), the parameter β value can be calculated in Eq. (12).

$$\beta = [H^T\ H]^{-1} H^T Y \tag{12}$$

In addition, β represents the output weights and Y shows real output.

2.4. Activation Function

In this study, four activation functions i.e. Scaled Exponential Linear Unit (SELU), Rectified Linear Unit (ReLU), Sigmoid, and Sine were used and compared for bearing degradation prediction [17], [18]. A brief description of each activation function is presented as follows:

2.4.1. Scaled Exponential Linear Unit

Scaled Exponential Linear Unit (SELU) is an activation function commonly used in Feed Forward Neural Network (FNN). The SELU activation function is typically performed satisfactorily on FNN networks. The SELU activation function is described in Eq. (13).

$$f(x) = \begin{cases} \alpha(e^z - 1), z < 0 \\ \lambda(z), z \geq 0 \end{cases} \tag{13}$$

where α is the slope at the negative part of the function and λ is scaling factor that functions to regulate the output amplitude, ensuring the network operates effectively without requiring additional normalization layers. In this study α of 1.5 and λ of 1.05 were used in the calculation.

2.4.2. Rectified Linear Unit (ReLU)

ReLU is the most widely used activation function among deep learning researchers. ReLU is linear for all positive values, and zero for all negative values. ReLU activation function is described in Eq. (14).

$$ReLU(x) = \max(0, x) \tag{14}$$

2.4.3. Sigmoid

Sigmoid is a mathematical function that converts inputs into outputs between 0 and 1, often used in artificial neural networks to generate probability values. The Sigmoid activation function is shown in Eq. (15).

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{15}$$

2.4.4. Sine

Sine is an activation function that uses the sine trigonometric function to convert inputs into outputs oscillating between 0 and 1. The Sine activation function is described in Eq. (16).

$$sin(x) \tag{16}$$

2.5. Evaluation Matrix

To assess the prediction performance of RVFL method for slew bearing prognosis, the evaluation metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) were used as presented in Eqs. (17) to (19) [19], [20].

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(X_i - Y_i)^2} \tag{17}$$

$$MAE = \frac{1}{m}\sum_{i=1}^{m}|X_i - Y_i| \tag{18}$$

$$MAPE = \frac{1}{m}\sum_{i=1}^{m}\left|\frac{X_i - Y_i}{X_i}\right| x\ 100\% \tag{19}$$

where $\boldsymbol{X_i}$ and $\boldsymbol{Y_i}$ is the actual and predicted value, respectively.

2.6. Comparable Methods

2.6.1. Extreme Learning Machine (ELM)

The single hidden layer feedforward networks, also known as SLFNs, are among the most well-known types of feedforward neural networks [21]. Their learning capabilities and fault tolerance properties have been widely discussed in both theoretical and practical studies. The recent development of the Extreme Learning Machine (ELM) neural algorithm for SLFNs has been utilized to improve their performance [21]. ELM, as a relatively new learning method for feedforward neural networks, differs from conventional neural networks in that the hidden biases and input weights are randomly initialized and remain fixed throughout the learning process. Additionally, the output weights are determined analytically. These characteristics contribute to ELM's strong generalization performance and fast learning speed [22].

Extreme learning machine (ELM) is a training algorithm for single hidden layer feedforward neural network (SLFN), which converges much faster than traditional methods and yields promising performance. Due to its exceptional performance, Extreme Learning Machine (ELM) has been widely utilized in various real-time learning applications, including classification, clustering, and regression tasks [23]. The ELM has the advantages of not falling into a local minimum easily and possessing stronger generalization ability than traditional methods, making it widely applicable in many fields [24].

The process of building an ELM model typically follows three sequential steps: (I) an SLFN is created; (II) weights and biases of the network are randomly selected; (III) the output weights are estimated by inverting the hidden layer output matrix [25]. In the parameter random initialization stage, the hidden layer parameters are randomly initialized, and the activation function is determined. The activation function is a nonlinear mapping that maps the input data to the ELM feature space. Specifically, parameter initialization randomly generates the weight (w) and bias (b) of the hidden layer nodes [26].

2.6.2. Artificial Neural Network

Artificial Neural Network (ANN) is one example of computational intelligence models and a good computational technique for modifying output based on input parameters and sparse experimental data [27]. ANNs are advanced computational frameworks that have evolved from the conceptual understanding of biological neural networks found in the human brain. These systems attempt to emulate human cognitive processes in a simplified, mathematical form. Although commonly associated with artificial intelligence, ANN predominantly operate on numerical and structured datasets. They present notable limitations when directly processing unstructured data types such as images, textual content, and audio signals [28].

The ANN consist of interconnected layers of artificial neurons that process data and learn patterns from input features. ANNs are widely used in regression and classification tasks due to their ability to model complex relationships in data. By adjusting the network's weights through training, ANNs can generalize patterns and make accurate predictions on unseen data.

A single-layer neural network is referred to as a perceptron. However, most practical implementations utilize multiple interconnected layers, forming what is known as a Multilayer Perceptron (MLP). These networks comprise numerous neurons or units, each connected to others in adjacent layers. To introduce non-linearity and enable the modeling of complex relationships within the data, various activation functions are employed across the layers. Common activation functions include the Sigmoid (Logistic), Tanh, ReLU, and Leaky ReLU functions [27]. A simplified neural network model can be represented mathematically as presented in Eq. (20) [29].

$$h_\theta(x) = \frac{1}{1 + e^{-\theta x^T}} \tag{20}$$

where $h_\theta(x)$ is the output, and *x* and θ are the parameter vectors input.

3. Experimental Setup and Data Acquisition

3.1. Low-speed Slew Bearing Test-rig

The experimental run-to-failure vibration bearing data in continuous mode were collected from a slew-bearing test rig adopted from steel mill industry. The test rig is designed to simulate a local steel-making company's actual working conditions, such as high applied load and very low rotational speed. The rotational speed of the slew-bearing test rig is adjustable from 1 to 12 rpm using an inverted motor controller and gear reducer mechanism. The slew-bearing test rig picture is presented in **Figure 2a**. A schematic of laboratory slew-bearing test rig showing a slew bearing attached in the drive ring, the applied load from the hydraulic, and other detail equipment labels is presented in **Figure 2b**.

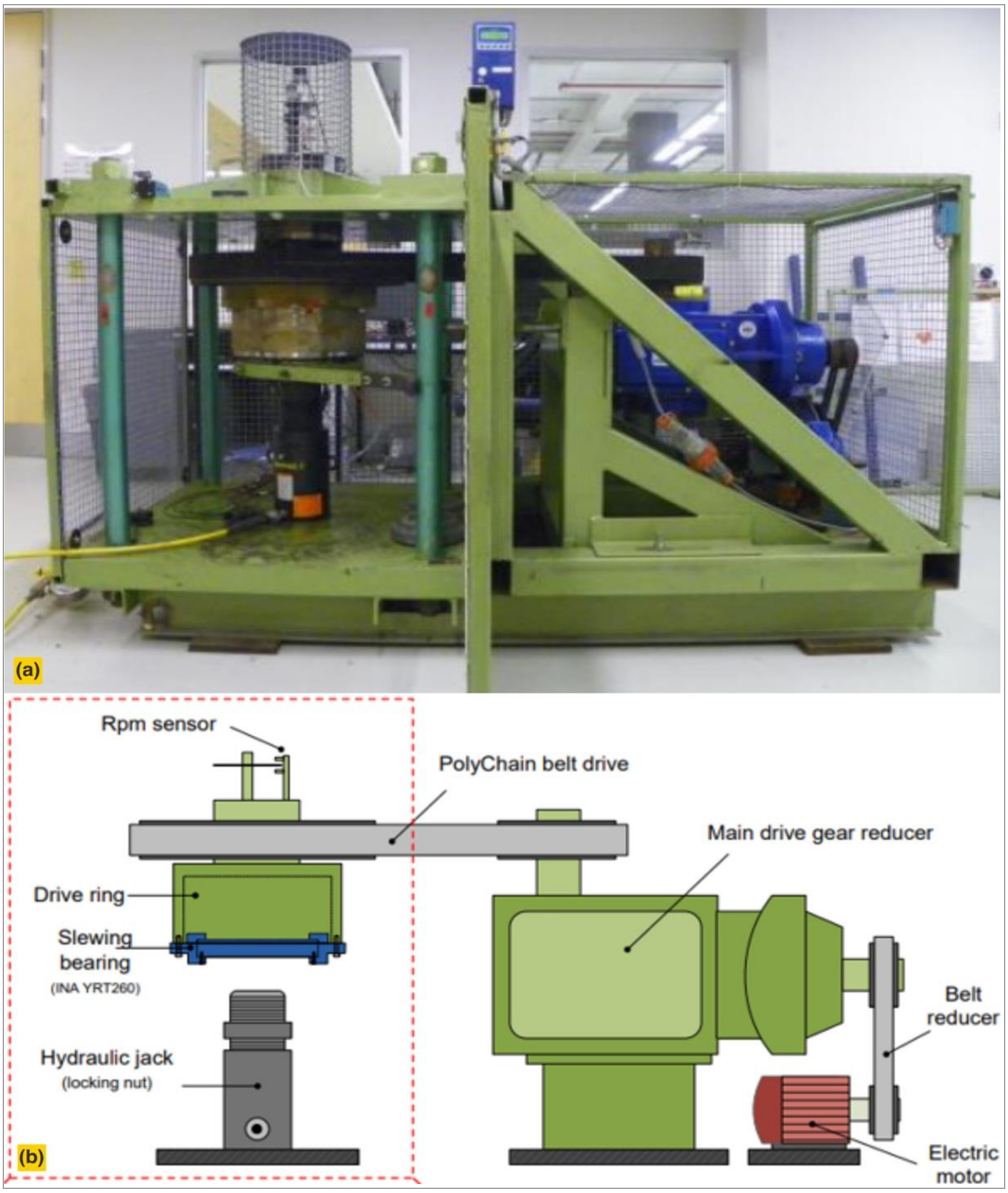


Figure 2. Slew bearing laboratory test rig: (a) Actual picture; (b) Schematic drawing

A low rotational speed variation was obtained using two systems: (1) a motor to main gear reducer double vee belt system, and (2) a final drive polychain belt drive. The test rig is designed comprehensively to consider safety and external vibration by using appropriate stress frame design specification and damper on the bottom side. The test rig consists of two control modes: (1) for continuous rotation and (2) for reversible rotation. The reversible rotation is a typical application in steel mining industry where the slew bearing run continuously in reversible rotation at 180 angle rotations.

3.2. Vibration Signal and Feature Extraction

The vibration data of low-speed slew bearing experiment was acquired from four accelerometers installed on the inner radial surface at 90° to each other with the sampling rates of 4880 Hz [2]. The accelerometers were IMI 608A11 ICP-type sensors. The accelerometers were connected to a high-speed Pico scope DAQ (PS3424). The bearing was subjected to an axial load of 15 tonnes. To provide continuous monitoring and produce run-to-failure bearing data, the bearing data was collected from February to August 2007 (139 days). In order to accelerate the bearing service life, coal dust was injected into the bearing in mid-April 2007 (58 days from the beginning) to simulate the actual working condition. In practice, especially in steel-making companies, the slew bearing is located in the open air and exposed to a dusty environment. A picture showing accelerometer sensors attachment on the slew bearing is presented in **Figure 3**. An example of the raw vibration signal on August 30 is presented in **Figure 4a**. The vibration data in time domain is transformed to frequency domain using Fast Fourier Transform (FFT) as presented in **Figure 4b**. It is shown in **Figure 4b** that the frequency content of the slew bearing is dominated with the low frequency spectrum.

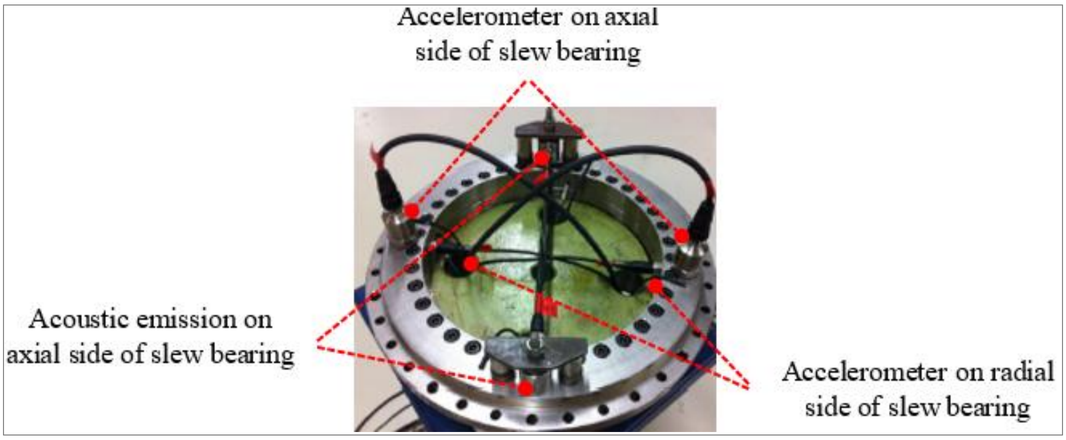


Figure 3. Acoustic emission and accelerometer sensor attached on the slew bearing

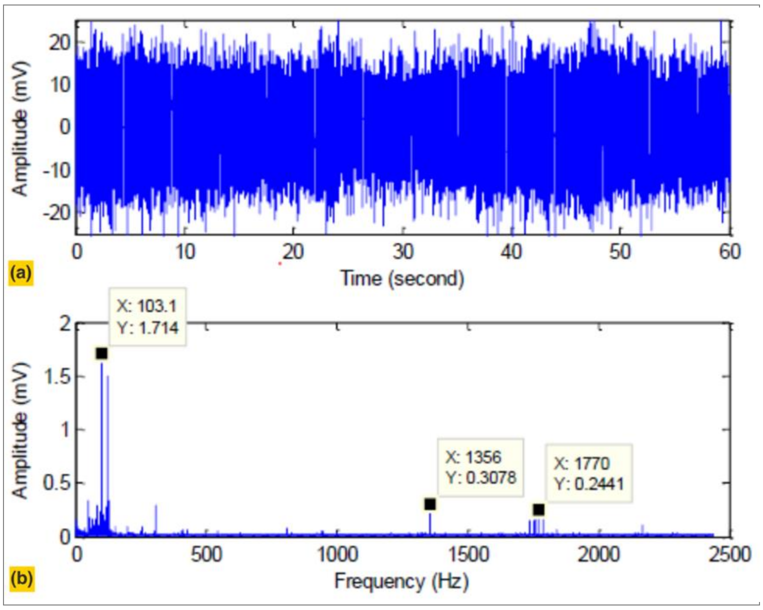


Figure 4. Vibration data acquired on 30 August: (a) Raw vibration data; (b) FFT

4. Result and Discussion

4.1. Block Diagram of Prognostics Method based on RVFL

A block diagram of proposed prognostics method based on RVFL method that applied in slew bearing vibration data is presented in **Figure 5**. The method starts with the feature extraction of the raw vibration data and followed by the degradation index calculation. The RVFL model is developed based on the training data and the build model is tested using the testing data. The RVFL model development include the selection of the activation function that is presented further in Section

4.4. The prediction result was evaluated using the evaluation matrix that presented in Section 2.5. If the evaluation matrix satisfied the criteria the next process is determine the remaining useful life (RUL). The highest accuracy of the activation function according to the evaluation matrix is selected as the best model.

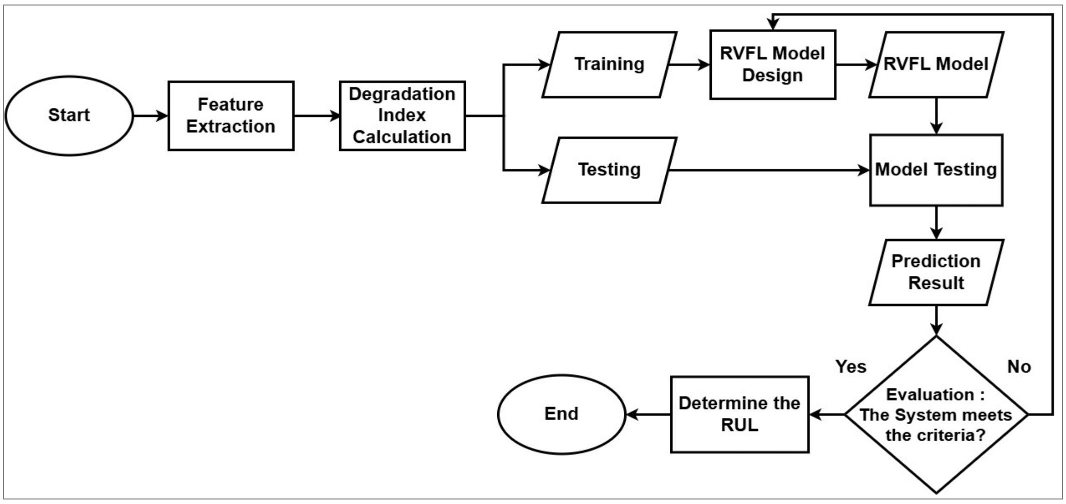


Figure 5. A block diagram of prognostic method for RUL prediction

4.2. Feature Extraction

There are 5 time-domain features were used in this study i.e., RMS, variance, kurtosis, histogram upper and histogram lower. A detail equation of these features has been presented in Section 2.1 and for further reading can be found in [2]. These features were calculated daily from the brand-new condition of bearing until bearing final failure with a total day of 139 days. The example plots of each feature is presented in **Figure 6a** to **Figure 6e**. The x-axis is the bearing operation days continuously from brand new until final failure (139 days); and the y-axis represents

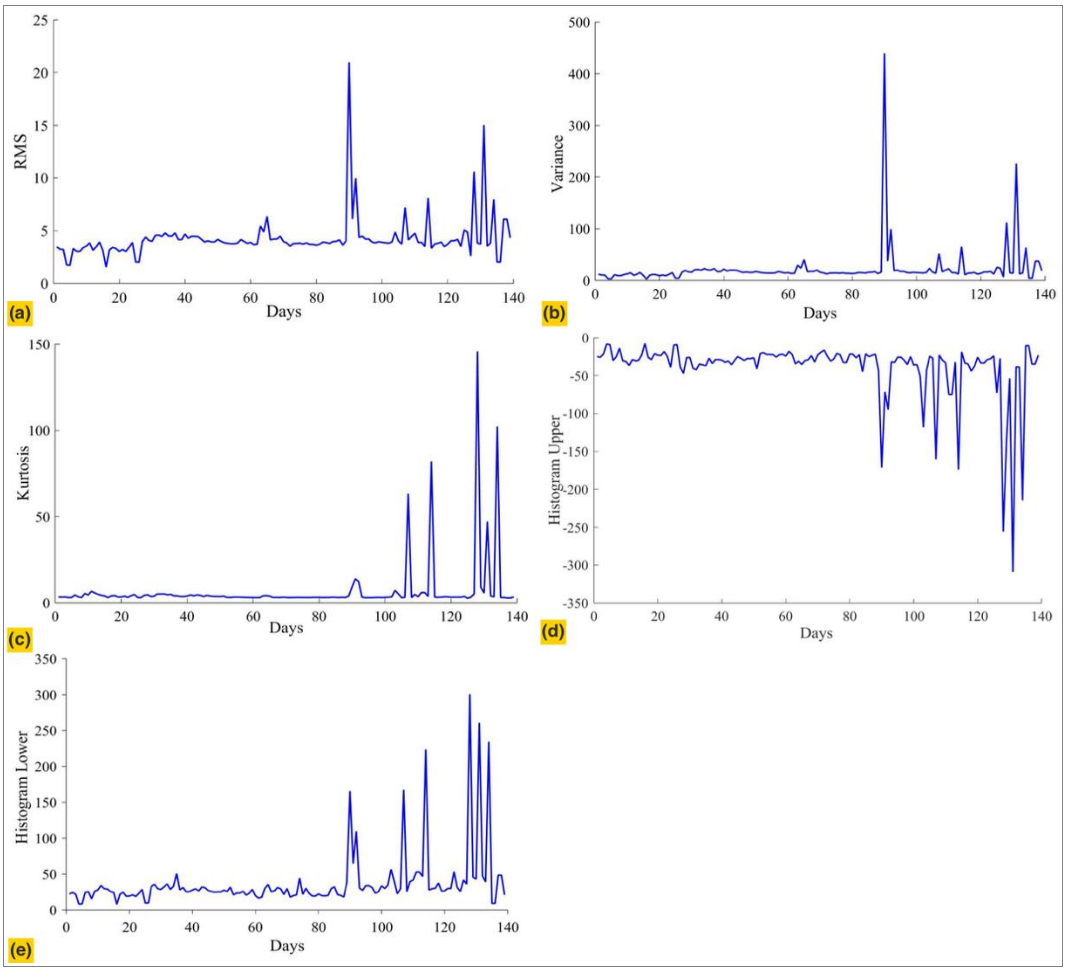


Figure 6. Time-domain features of slew bearing vibration data in 139 days: (a) Root mean square; (b) Variance; (c) Kurtosis; (d) Histogram upper; (e) Histogram lower

the value of each feature. The features data were processed further to obtain the degradation index using the formula presented in Section 2.2. For the RVFL prediction, the features data in the form of the degradation index were also divided into two datasets, i.e., (1) for training, and (2) for testing. A detail explanation is presented in Section 4.5.

According to **Figure 6a** to **Figure 6e**, five features show a low amplitude level and low fluctuation from day 1 to day 90. This indicate that the raw vibration signal is still within the acceptable range and it is associated to the normal bearing condition. A sudden changed was obviously appeared at day 90 where five features show an initial peak amplitude. This condition is corresponded to the initial failure of the bearing condition. Moving forward from day 90, the features were showing the increasing amplitude level until the bearing had a final failure at 139 days. At 139 days the bearing is totally stop due to harsh noise sound and the highest vibration amplitude. The bearing is then dismantled and it was found that the inner and outer race as well as the rollers were in harsh crack and spall condition. The picture of these condition can be found in [30].

4.3. Degradation Index

After features were calculated, the index Degradation Index (DI) is determined and the actual value at the peak of the index degradation value can be obtained. **Figure 7** shows the actual value at the peak of the index degradation value. The degradation index in **Figure 7** was calculated using Eq. (6). The initial failure threshold of the degradation index was set to 90 as this number is almost a half of the highest value of DI. Later, these DI values will be predicted using RVFL and other two prediction methods (Extreme Learning Machine (ELM) and Artificial Neural Network (ANN)) to obtained a RUL prediction. **Table 1** is a detail values of the DI that will be used as a reference or an actual value for a RUL prediction and error estimation.

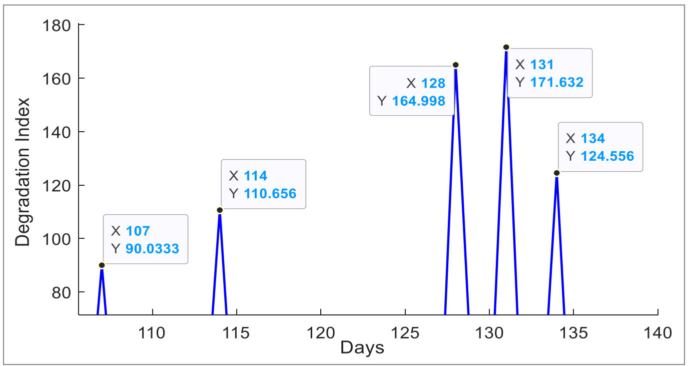


Figure 7. Actual degradation Index based on five time-domain features

Table 1. Actual degradation index value (calculated from five time-domain features)

Day	Degradation Index
107	90.03
114	110.656
128	164.998
131	171.632
134	124.556

4.4. Activation Functions

In this paper, different activation functions were used to show whether the results obtained were significantly different or not. In this study, a ratio data 80:20 is used with 80% for training data and 20% for test data and activation function used. The selected activation function for this study are SELU, ReLU, Sigmoid, and Sine. In each activation function, 30 trials were performed to find the minimum RMSE. The 30 trials result of four selected activation function is presented in **Figure 8**. In general, the lower RMSE value was obtained in **Figure 8a** and **Figure 8b** for SELU and ReLU compare to **Figure 8c** and **Figure 8d** for Sigmoid and Sine. The average RMSE value of SELU, ReLU, Sigmoid, and Sine is 5.8427, 8.5513, 23.6733, and 83.801, respectively. Among 30 trials, the lowest RMSE was noted and used as the prediction model for RUL bearing prognostic. The lowest RMSE value for SELU. ReLU, Sigmoid, and Sine of 0.9598, 1.9643, 11.1054, and 48.8755, respectively. A general activation function comparison is presented in **Table 2**. For a more detail comparison including the optimum trial number and the evaluation metrics is presented in **Table 3**. In order to provide a comprehensive comparison for the model that used four different activation function, the degradation index prediction for four activation function is presented in **Figure 9**.

Table 2. Activation function comparison (in general)

Activation function	Average of RMSE for 30 trials	The lowest RMSE	The highest RMSE
SELU	5.6769	0.9598	23.3402
ReLU	8.5515	1.9643	31.0234
Sigmoid	23.6734	11.1054	43.7997
Sine	83.7007	48.8755	213.977

Table 3. Activation function comparison (for the lowest RMSE, MAE and MAPE)

Activation function	The lowest trial number	The lowest RMSE	The lowest MAE	The lowest MAPE (%)
SELU	25	0.9598	0.4474	0.9904
ReLU	12	1.9643	0.9102	2.29
Sigmoid	20	11.1054	5.4159	12.8199
Sine	4	48.8755	26.8544	56.1383

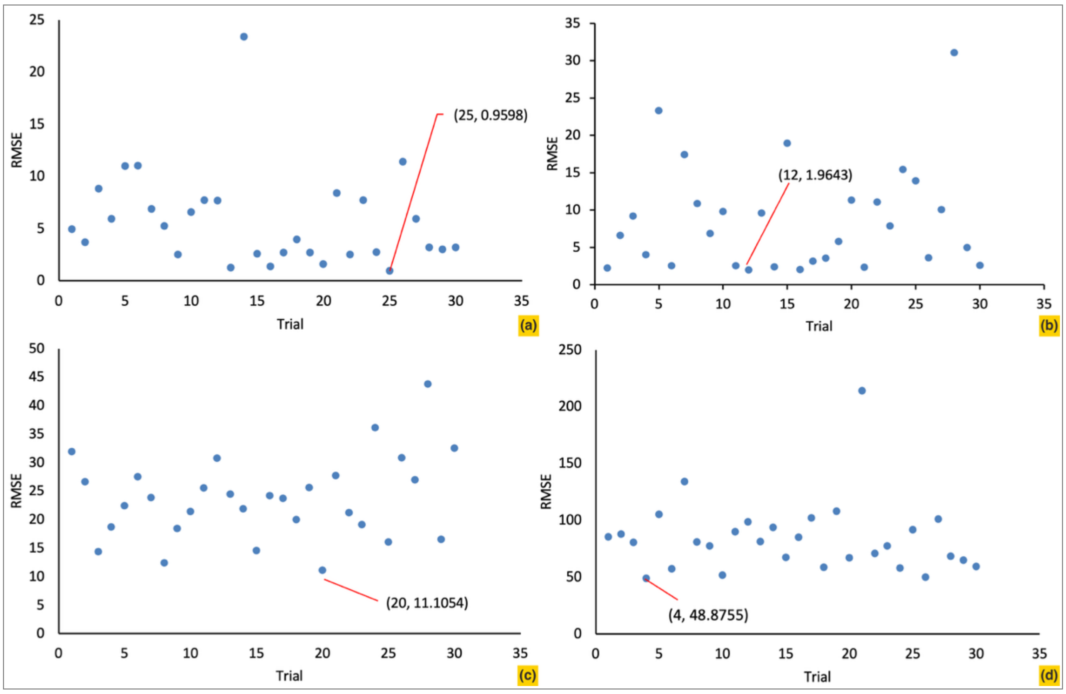


Figure 8. RMSE of four activation functions for 30 trials: (a) SELU; (b) ReLU; (c) Sigmoid; (d) Sine

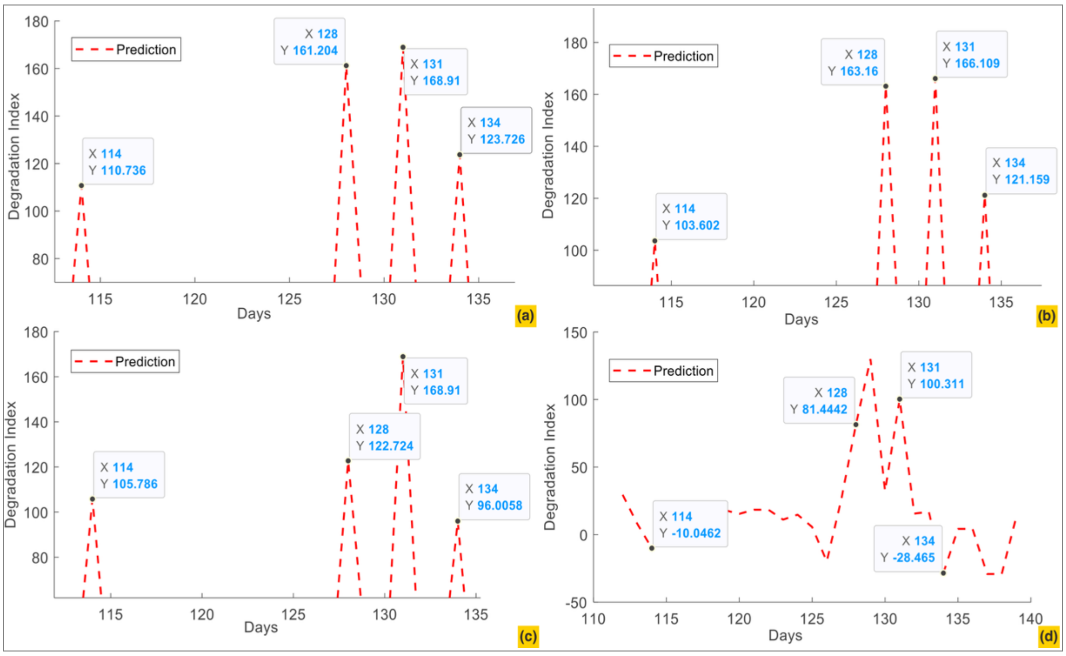


Figure 9. Degradation index prediction: (a) SELU; (b) ReLU; (c) Sigmoid; (d) Sine

According to activation function comparison results presented in **Figure 8** and **Table 3**, the SELU activation function at trial number 25 was selected as the prediction model for the degradation index.

4.5. Remaining Useful Life (RUL) or RVFL method

To determine the RUL value of the slew bearing, the best activation function is selected. According to the comparison of the activation functions, SELU is outperformed that other activation function. Then to determine the threshold value, the first time the bearing is damaged is used, namely on the 90th day. The threshold values obtained are as follows:

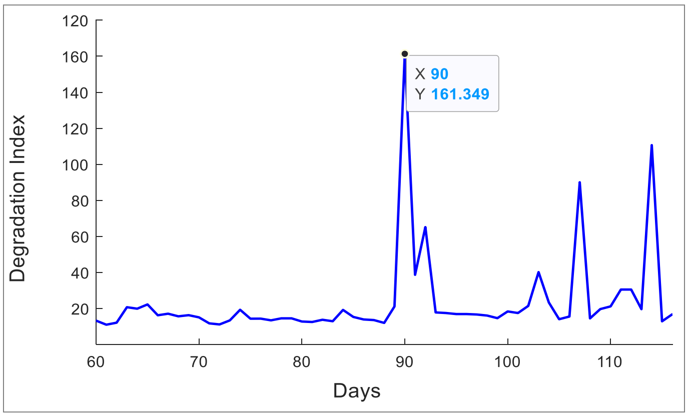


Figure 10. Peak of degradation index

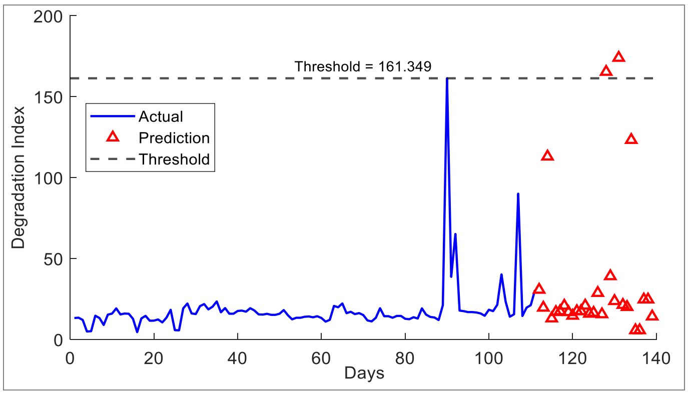


Figure 11. RVFL prediction

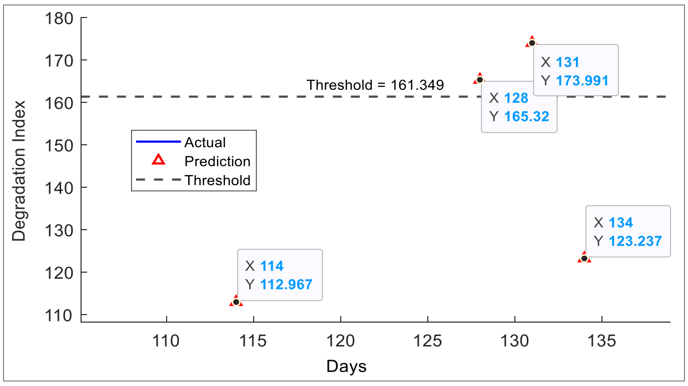


Figure 12. RUL prediction (zoom-in view of the red dashed line of **Figure 11**)

As presented in **Figure 10**, the peak point value on day 90 is 161.349 (≈ 161). This value was used as a final failure threshold of the DI is determined for the RUL prediction of the bearing. This threshold indicates the level of the bearing condition near before the final failure occurred. Then after obtaining the threshold value, the threshold value is entered into the experiment using the best activation function.

Figure 11 shows the RUL bearing prediction obtained from the developed model, RVFL. From the figure, it is shown that the RUL prediction stops on day 131 because the point stops above the threshold value. This indicates that the bearing will be damaged on day 131. However, based on the actual data obtained, the bearing was damaged on day 139. The difference obtained from the actual data value and the prediction data is 8 days. From this difference, the accuracy value of the RUL prediction is calculated. The zoom in view (the red dashed line) of **Figure 11** is presented in **Figure 12**.

To calculate the accuracy of the RUL prediction, the following formula is used [2].

$$PA = \left(1 - \frac{|Actual\ RUL - Predicted\ RUL|}{Actual\ RUL}\right) x\ 100\% \tag{21}$$

where, PA is a prediction accuracy of the RUL prediction.

By substituting the actual and predicted RUL to Eq. (20), the RUL prediction accuracy can be obtained as follows:

$$PA = \left(1 - \frac{|139 - 131|}{139}\right) x\ 100\%$$

$$PA = \left(1 - \frac{|8|}{139}\right) x\ 100\% \tag{22}$$

$$PA = 0.9424\ x\ 100\%$$

$$PA = 94.24\%$$

According to the above calculation, the RUL prediction accuracy of RVFL method is 94.24%. The RUL value obtained from the above calculation is acceptable because the accuracy value is above 90%. This also indicates that the model used is good enough to predict when the bearing will be damaged.

The RVFL method has a lower accuracy value compared to the existing method (Kernel Regression), namely because the RVFL method gets random results and requires several trials to get maximum results. So that with random results and the selection of the best activation function, the accuracy obtained meets the design criteria but does not exceed or lower than the existing methods.

Suggestion to improve accuracy with RVFL are changing the alpha and lambda values in the SELU activation function, however, this method will be time consuming because it is a trial-and-error method. Another alternative is by combining the RVFL method with other ML methods to optimize the alpha and lambda values.

4.6. Results from Extreme Learning Machine and Artificial Neural Network

The following are the comparison result of the RVFL method with ELM and ANN methods.

4.6.1. ELM Prediction

The DI prediction result of ELM method is presented in **Figure 13**. Similar to the RVFL prediction, four DI points (day 114, 128, 131, and 134) were also predicted using ELM and it is shown in detail in **Figure 13**. It is observed that when utilizing the ELM, the highest point is reached on day 128. Using the RUL prediction accuracy shows in Eq. 20 the prediction accuracy (PA) of ELM is 92.08%. This PA of ELM is lower than the PA of RVFL. This prediction indicated that the RUL estimation is earlier than both the RVFL prediction and actual value. Other detail comparison based on the DI values of the four days prediction (day 114, 128, 131, and 134) is presented in **Table 4** to **Table 6**.

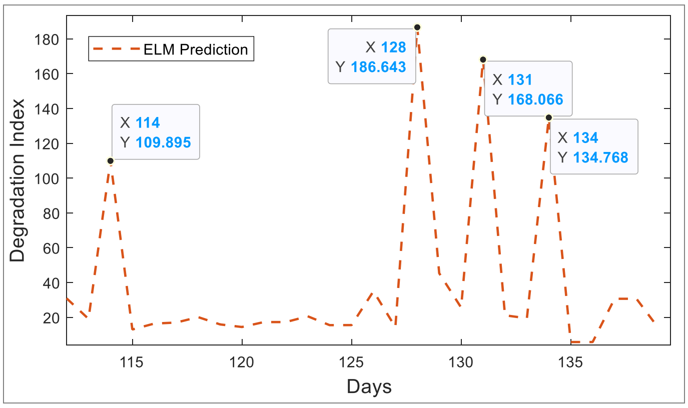


Figure 13. ELM prediction

4.6.2. ANN (Artificial Neural Network)

The DI prediction of four DI points (day 114, 128, 131, and 134) using ANN method is presented in **Figure 14**. Similar to ELM prediction, the highest prediction point was on day 114. This will result to a similar RUL prediction accuracy of 92.08%. The ANN RUL prediction also earlier than the RVFL RUL prediction. Other detail comparison based on the DI values of the four days prediction is presented in **Table 4** to **Table 6**.

Table 4 presents the prediction accuracy of RUL for only the highest peak day. The four prediction days of DI values is presented in **Table 5** compare to the actual DI value. In **Table 6**, the prediction is presented in percentage to inform the best performance out of three methods.

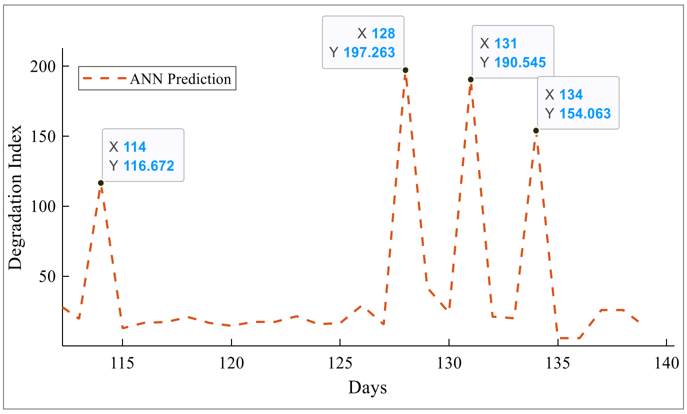


Figure 14. ANN prediction

Table 4. Comparison of RUL prediction using RVFL, ELM, and ANN

Method	Peak day	Prediction Accuracy of RUL (%)
RVFL	131	94.24
ELM	128	92.08
ANN	128	92.08

Table 5. Comparison of DI level prediction using RVFL, ELM, and ANN

Day of Predicted DI	Actual DI value	RVFL DI Prediction	ELM DI Prediction	ANN DI Prediction
114	110.656	112.967	109.895	116.672
128	164.998	165.32	186.643	197.263
131	171.632	173.991	168.006	190.545
134	124.556	123.237	134.768	154.063

Table 6. Comparison of DI level prediction using RVFL, ELM, and ANN in percentage

Day of Predicted DI	RVFL DI Prediction in %	ELM DI Prediction in %	ANN DI Prediction in %
114	97.91	99.31	94.56
128	99.80	86.88	80.45
131	98.63	97.89	88.98
134	98.94	91.80	76.31
Average	98.82	93.97	85.07

5. Conclusions

The RVFL model for obtaining the RUL of bearings has met the design criteria with an accuracy of 94.24%. The peak value prediction obtained using RVFL with an 80:20 data ratio and the SELU activation function has met the design criteria with evaluation metrics of RMSE at 0.9598, MAE at 0.4474, and MAPE at 0.9904%. Furthermore, the difference between the actual values and the predicted values is minimal. The recommendations provided are to modify the parameter values in the SELU activation function, use a larger training dataset, and combine the RVFL method with other ML techniques.

As the data is limited that only used one bearing, the future works is adding a number of bearings as the research object and change the parameters such as applied load, rotating speed, and day of dust insertion. The greater number of bearings used in the bearing condition monitoring and prognosis research with differs operating parameters will represent closely on the actual condition.

Acknowledgement

The authors thank to Diponegoro University for the Adjunct Professor program No. 1276/UN7.J7/DK/2024.

Authors' Declaration

Authors' contributions and responsibilities - The authors made substantial contributions to the conception and design of the study. The authors took responsibility for data analysis, interpretation, and discussion of results. The authors read and approved the final manuscript.

Funding – This study is funded by Diponegoro University from the Adjunct Professor program No. 1276/UN7.J7/DK/2024.

Availability of data and materials - All data is available from the authors.

Competing interests - The authors declare no competing interests.

Additional information – No additional information from the authors.

References

[1] F. Xu *et al.*, "A review of bearing failure Modes, mechanisms and causes," *Engineering Failure Analysis*, vol. 152, p. 107518, Oct. 2023, doi: 10.1016/j.engfailanal.2023.107518.

[2] W. Caesarendra, "Vibration and acoustic emission-based condition monitoring and prognostic methods for very low speed slew bearing," University of Wollongong, 2015.

[3] G. V. P. Babu and M. P. Devi, "Slew Bearing Back Lash Checking Importance for Stackers/ Reclaimers and Mobile Harbour Cranes for the Availability of the Equipment-Ports Service Quality," *TEST Engineering and Management*, 2020.

[4] T. Gao, Y. Li, X. Huang, and C. Wang, "Data-Driven Method for Predicting Remaining Useful Life of Bearing Based on Bayesian Theory," *Sensors*, vol. 21, no. 1, p. 182, Dec. 2020, doi: 10.3390/s21010182.

[5] R. N. Toma, A. E. Prosvirin, and J.-M. Kim, "Bearing Fault Diagnosis of Induction Motors Using a Genetic Algorithm and Machine Learning Classifiers," *Sensors*, vol. 20, no. 7, p. 1884, Mar. 2020, doi: 10.3390/s20071884.

[6] D. Maulud and A. M. Abdulazeez, "A Review on Linear Regression Comprehensive in Machine Learning," *Journal of Applied Science and Technology Trends*, vol. 1, no. 2, pp. 140–147, Dec. 2020, doi: 10.38094/jastt1457.

[7] S. Zhang, S. Zhang, B. Wang, and T. G. Habetler, “Deep Learning Algorithms for Bearing Fault Diagnostics—A Comprehensive Review,” *IEEE Access*, vol. 8, pp. 29857–29881, 2020, doi: 10.1109/ACCESS.2020.2972859.

[8] A. Z. Hinchi and M. Tkiouat, “Rolling element bearing remaining useful life estimation based on a convolutional long-short-term memory network,” *Procedia Computer Science*, vol. 127, pp. 123–132, 2018, doi: 10.1016/j.procs.2018.01.106.

[9] X. Chen, B. Zhang, and D. Gao, “Bearing fault diagnosis base on multi-scale CNN and LSTM model,” *Journal of Intelligent Manufacturing*, vol. 32, no. 4, pp. 971–987, Apr. 2021, doi: 10.1007/s10845-020-01600-2.

[10] R. Katuwal, P. N. Suganthan, and M. Tanveer, “Random Vector Functional Link Neural Network based Ensemble Deep Learning,” *Computer Science*, Jun. 2019, [Online]. Available: http://arxiv.org/abs/1907.00350

[11] A. K. Malik, R. Gao, M. A. Ganaie, M. Tanveer, and P. N. Suganthan, “Random vector functional link network: Recent developments, applications, and future directions,” *Applied Soft Computing*, vol. 143, p. 110377, Aug. 2023, doi: 10.1016/j.asoc.2023.110377.

[12] W. Dai, Q. Chen, F. Chu, X. Ma, and T. Chai, “Robust Regularized Random Vector Functional Link Network and Its Industrial Application,” *IEEE Access*, vol. 5, pp. 16162–16172, 2017, doi: 10.1109/ACCESS.2017.2737459.

[13] S. Wang, J. Chen, H. Wang, and D. Zhang, “Degradation evaluation of slewing bearing using HMM and improved GRU,” *Measurement*, vol. 146, pp. 385–395, Nov. 2019, doi: 10.1016/j.measurement.2019.06.038.

[14] B. Y. Kosasih, W. Caesarendra, K. Tieu, A. Widodo, C. A. S. Moodie, and A. K. Tieu, “Degradation Trend Estimation and Prognosis of Large Low Speed Slewing Bearing Lifetime,” *Applied Mechanics and Materials*, vol. 493, pp. 343–348, Jan. 2014, doi: 10.4028/www.scientific.net/AMM.493.343.

[15] W. Cao, J. Gao, Z. Ming, S. Cai, and H. Zheng, “Impact of Probability Distribution Selection on RVFL Performance,” in *Lecture Notes in Computer Science*, 2018, pp. 114–124. doi: 10.1007/978-3-319-73830-7_12.

[16] S. P. Mishra *et al.*, “An efficient Robust Random Vector Functional Link network for Solar Irradiance, Power and Wind speed prediction,” in *2021 1st Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology(ODICON)*, IEEE, Jan. 2021, pp. 1–7. doi: 10.1109/ODICON50556.2021.9428977.

[17] R. Alkhatib, “Artificial Neural Network Activation Functions in Exact Analytical Form (Heaviside, ReLU, PReLU, ELU, SELU, ELiSH),” Aug. 06, 2021. doi: 10.36227/techrxiv.15096888.v1.

[18] A. D. Rasamoelina, F. Adjailia, and P. Sincak, “A Review of Activation Function for Artificial Neural Network,” in *2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, IEEE, Jan. 2020, pp. 281–286. doi: 10.1109/SAMI48414.2020.9108717.

[19] A. S. Ahmar, “Forecast Error Calculation with Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE),” *JINAV: Journal of Information and Visualization*, vol. 1, no. 2, pp. 94–96, Dec. 2020, doi: 10.35877/454RI.jinav303.

[20] T. O. Hodson, “Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not,” *Geoscientific Model Development*, vol. 15, no. 14, pp. 5481–5487, Jul. 2022, doi: 10.5194/gmd-15-5481-2022.

[21] D. C. E. Saputra, K. Sunat, and T. Ratnaningsih, “A New Artificial Intelligence Approach Using Extreme Learning Machine as the Potentially Effective Model to Predict and Analyze the Diagnosis of Anemia,” *Healthcare*, vol. 11, no. 5, p. 697, Feb. 2023, doi: 10.3390/healthcare11050697.

[22] M. Shariati, D. J. Armaghani, M. Khandelwal, J. Zhou, and M. Khorami, “Assessment of Longstanding Effects of Fly Ash and Silica Fume on the Compressive Strength of Concrete Using Extreme Learning Machine and Artificial Neural Network,” *Journal of Advanced Engineering and Computation*, vol. 5, no. 1, p. 50, Mar. 2021, doi: 10.25073/jaec.202151.308.

[23] J. Wang, S. Lu, S.-H. Wang, and Y.-D. Zhang, “A review on extreme learning machine,” *Multimedia Tools and Applications*, vol. 81, no. 29, pp. 41611–41660, Dec. 2022, doi: 10.1007/s11042-021-11007-7.

[24] H. Wei, Q. Zhang, M. Shang, and Y. Gu, “Extreme learning Machine-based classifier for fault diagnosis of rotating Machinery using a residual network and continuous wavelet transform,” *Measurement*, vol. 183, p. 109864, Oct. 2021, doi: 10.1016/j.measurement.2021.109864.

[25] M. Shariati *et al.*, “A novel hybrid extreme learning machine–grey wolf optimizer (ELM-GWO) model to predict compressive strength of concrete with partial replacements for cement,” *Engineering with Computers*, vol. 38, no. 1, pp. 757–779, Feb. 2022, doi: 10.1007/s00366-020-01081-0.

[26] Z. Yu, J. Liu, M. Yang, Y. Cheng, J. Hu, and X. Li, “An Elderly Fall Detection Method Based on Federated Learning and Extreme Learning Machine (Fed-ELM),” *IEEE Access*, vol. 10, pp. 130816–130824, 2022, doi: 10.1109/ACCESS.2022.3229044.

[27] P. Kulkarni, V. Watwe, T. Chavan, and S. Kulkarni, “Artificial Neural Networking for remediation of methylene blue dye using Fuller’s earth clay,” *Current Research in Green and Sustainable Chemistry*, vol. 4, p. 100131, 2021, doi: 10.1016/j.crgsc.2021.100131.

[28] G. D.R., “Application of Artificial Neural Networking Technique to Predict the Geotechnical Aspects of Expansive Soil: A Review,” *International Journal of Engineering and Manufacturing*, vol. 11, no. 6, pp. 48–53, Dec. 2021, doi: 10.5815/ijem.2021.06.05.

[29] D. A. Otchere, T. O. Arbi Ganat, R. Gholami, and S. Ridha, “Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models,” *Journal of Petroleum Science and Engineering*, vol. 200, p. 108182, May 2021, doi: 10.1016/j.petrol.2020.108182.

[30] W. Caesarendra, P. B. Kosasih, A. K. Tieu, C. A. S. Moodie, and B.-K. Choi, “Condition monitoring of naturally damaged slow speed slewing bearing based on ensemble empirical mode decomposition,” *Journal of Mechanical Science and Technology*, vol. 27, no. 8, pp. 2253–2262, Aug. 2013, doi: 10.1007/s12206-013-0608-7.