Journal of Engineering Science and Technology
Vol. 13, No. 1 (2018) 102 - 121

A STUDY OF VARIOUS BLADE FAULT CONDITIONS ON A WIND TURBINE USING VIBRATION SIGNALS THROUGH HISTOGRAM FEATURES

A. JOSHUVA, V. SUGUMARAN*

School of Mechanical and Building Sciences (SMBS), VIT University,
Chennai Campus, Vandalur-Kelambakkam Road, Chennai 600127, India
*Corresponding Author: v_sugu@yahoo.com

Abstract

Wind energy is one of the rapidly evolving renewable energy resources. It is particularly essential to built dependability and accessibility of wind turbines and additional to decrease the wind energy cost. Wind turbine blades are ought to be an important component among the other basic segments in the wind turbine framework since they transform the dynamic energy of wind into useable power. Wind turbine blades are manufactured from either carbon fiber reinforced polymer or glass fiber reinforced polymer. Damages and flaws are unavoidable either in the manufacturing process or the lifetime of a composite blade. Hence, structural health monitoring for wind turbine blade is essential to avoid failures and extend dependability in both fabrication quality control and in-service investigation. In this study, a three-bladed variable wind turbine was chosen and using histogram features, the condition of a wind turbine blade is inspected. The faults like hub-blade loose connection, blade crack, pitch angle twist, erosion and blade bend faults were studied and these faults are classified using various data mining algorithms. The main contribution of this study is to build and suggest a data-model for fault identification on wind turbine blade while in operation using machine learning classifiers like sequential minimal optimization (SMO) algorithm, simple logistic algorithm (SLA), multilayer perceptron (MLP), logistic algorithm (LA) and radial basis function (RBF).

Keywords: Structural health monitoring, Fault diagnosis, Wind turbine blade, Machine learning, Histogram features, Vibration signals.

1. Introduction

The wind energy is one of the rapidly growing renewable energy resources, and it will have been a remarkable place in the energy market in the upcoming era.

Nomenclatures

A	Related features
a	Slope
α_i	Lagrange Multipliers
b	Intercept
c	The number of classes
k	Negative of the sum over
P_i	The proportion of S belonging to class 'i'.
S	Samples or examples
S_v	The subset of S
$X1$ to $X100$	Bin value range
x_1, y_1	Dataset

Greek Symbols

β	Independent Variable
φ	Activation function
ω	Vector of weights
$\phi (x)$	Distance from the origin

Abbreviations

AANN	Auto-Associative Neural Network
ADC	Analogue-To-Digital Converter
ANN	Artificial Neural Network
BEM	Blade Element Momentum
BB	Blade Bend
BC	Blade Crack
BE	Blade Erosion
BG	Blade Good
BPT/PAT	Blade Pitch Twist/Pitch Angle Twist
CFD	Computational Fluid Dynamics
DAQ	Data Acquisition System
FP	False Positive Rate
HAWT	Horizontal Axis Wind Turbine
LA	Logistic Algorithm
LabVIEW	Laboratory Virtual Instrument Engineering Workbench
MLP	Multilayer Perceptron
NDT	Non-Destructive Testing
NI	National Instruments
NLPCA	Nonlinear Principal Component Analysis
O&M	Operations And Maintenance
PCA	Principal Component Analysis
RBF	Radial Basis Functions
ROC	Receiver Operating Characteristics
SLA	Simple Logistic Algorithm
SMO	Sequential Minimal Optimization
TP	True Positive Rate
WEKA	Waikato Environment For Knowledge Analysis

As per gathered statistical information, it is projected that wind energy can supply around 12% of the worldwide power supply by 2020, and transcend 20% by 2030 [1]. As the wind energy region develops, business financial aspects will request gradually vigilant administration of costs [2]. The operations and maintenance (O&M) expenses of wind turbines are around 25-30% of the complete energy production cost [3]. Keeping in mind, to decrease wind energy costing, the decrement of the operations and maintenance expense is must [4]. Aside from applying enhancement design for a machine to develop the accessibility, another simple way is utilizing structural health monitoring for wind turbines.

The wind turbine is a distinctive Mechatronics structure. It is comprised of various mechanical and electrical segments, including the generator, blades, bearings, rotor, gearbox, tower, shaft, pitch and yaw [5]. Among these segments, blades are considered as a critical and significant segment [6]. Since the efficiency of wind turbine captures the wind energy depends upon the propeller-like blades. Moreover, the blade fabricating expense is nearly 15-20% of each wind turbine cost. Some flaws occurred during the fabrication process and some faults are caused due to environmental conditions. If the blade damage is large then it will cause terrible destruction to the surroundings and may also damage the whole turbine structure [7]. It will create a huge loss in power production. So to reduce the loss of productivity, condition monitoring is preferred to find the damage while the turbine is in operating condition.

There are two types of approaches which are carried out for condition monitoring: traditional approach and machine learning approach. The traditional approach is mainly used where frequency component does not change with respect to time. Rotating machines produce non-stationary signals. Since the frequency components change due to wear and tear, fault discrimination is very difficult using an automated system in the traditional approach. Hence, it is not preferred. In machine learning approach, algorithms have the capability to learn continuously and adapt themselves to the varying situations. So researchers often resort to machine learning approach for fault diagnosis of mechanical systems.

Many studies and research are carried out in condition monitoring of wind turbine blades, to mention a few. Frost et al. [8] carried out a study on integrating structural health management with contingency control for wind turbines using nonlinear high-fidelity simulation. The structural health and contingency control of the blade was studied. The speed of the turbine and decision making using prognostic information was also carried out. A study on damages of wind turbine blade trailing edge (forms, location, and root causes) was carried out by Ataya and Ahmed [9]. This paper analyzed and studied about the crack location on wind turbine blades (both longitudinal cracks and transverse cracks) using Non-Destructive Testing (NDT) method and discussed the life of the blade.

A work on damage diagnosis for a wind turbine blade using pattern recognition was carried out by Dervilis et al. [10]. This study carried out the condition monitoring of the blade using algorithms like principal component analysis (PCA), nonlinear principal component analysis (NLPCA), artificial neural network (ANN), auto-associative neural network (AANN), and radial basis functions (RBF). Roth-Johnson et al. [11] carried out a structural design of spars for 100 m biplane wind turbine blades by beam finite elements with a cross-sectional analysis. This paper mainly focuses on the wind turbine blade design.

Lee et al. [12], done a work on wind turbine blade moment signals to blade condition monitoring using a transformation algorithm. This study presented a novel method of transforming blade moment signals on a horizontal axis 3-blade wind turbine. A validation of an integral sliding mode control for optimal control of a three blade variable speed variable pitch wind turbine was done by Saravanakumar and Jena [13]. This study mainly focuses on the control of variable speed variable pitch wind turbine for maximization of extracting power at below rated wind speed and regulation of extracting power when operating at above-rated wind speed.

Vučina et al. [14] have done a numerical model for robust shape optimization of wind turbine blades using 3D geometric modeler. A computational framework for the shape optimization of wind turbine blades is developed for variable operating conditions specified by local wind speed distributions. This study considered the blade design using simulation process, and it didn't focus on the faults which affect the performance of the wind turbine. Aero-structural design and optimization of a small wind turbine blade study were carried out by Pourrajabian et al. [15]. This study developed a methodology for aero-structural design, including consideration of the starting of a small wind turbine blade. This study carried out for both structural analysis and stress analysis on the blade by optimization.

Bessa et al. [16] carried out a work on data-driven fault detection and isolation scheme for a wind turbine benchmark. In their first step, the fault detection is based on an alternative method based on the Gibbs sampling algorithm in which the occurrence of a sensor fault is modeled as a change point detection in a time series. On the second step, the fault isolation is handled via Fuzzy/Bayesian network scheme classifying the kind of fault. Simulation of aeroelastic behavior in a composite wind turbine blade was studied by Rafiee et al. [17]. Aero-elastic analysis of a full-scale composite wind turbine blade was investigated using 3D model and aerodynamic loading was determined using modified Blade Element Momentum (BEM) theory and Computational Fluid Dynamics (CFD) method.

Optimal selection of autoregressive model coefficients for early damage detects ability with an application to wind turbine blades was carried out by Hoell and Omenzetter [18]. This study enhances the selection of autoregressive model coefficients for statistical hypothesis testing for damage presence and adding or eliminating the coefficients is carried out by genetic algorithm. Rezaei et al. [19] carried out a study on modal based damage identification for the nonlinear model of modern wind turbine blade. This study considered geometric nonlinearity due to the large structural deformation of the modern wind turbine blade using a finite element model.

Numerous works were carried out using simulation analysis of fault and design analysis of wind turbine blade; however, only a very few in the experimental analysis was carried out. The machine learning technique was considered for wind turbine blade fault diagnosis; however, the usage was limited in the literature. A very limited set of defects was considered for analysis. This is especially true in the case of fault diagnosis of wind turbine blade. Hence, there is a strong need to design a fault diagnosis system which can handle multiple faults in wind turbine blades using machine learning approaches. This study makes a novel attempt to find different blade faults applying machine learning approach and histogram analysis. Figure 1 shows the flowchart of the work done. The contribution of the present study,

- This study considers five faults (blade crack, erosion, hub-blade loose connection, pitch angle twist and blade bend) for wind turbine blade fault diagnosis.
- Histogram feature extraction tool was used to extract the required features from the vibration signals.
- J48 decision tree algorithm was used for feature selection.
- This problem is modeled as a multiclass classification problem and attempts to classify using machine learning classifiers like sequential minimal optimization (SMO) algorithm, the simple logistic algorithm (SLA), multilayer perceptron (MLP), the logistic algorithm (LA) and radial basis function (RBF).

The rest of the paper is organized as follows. Section 2 presents the experimental setup and experimental procedure is explained. In section 3, feature extraction is explained, followed by feature selection in section 4. The classifiers used in this study are explained in section 5. The classification accuracy of the models was discussed and the suggestion of the better model is proposed in section 6. Conclusions are presented in the final section (section 7).

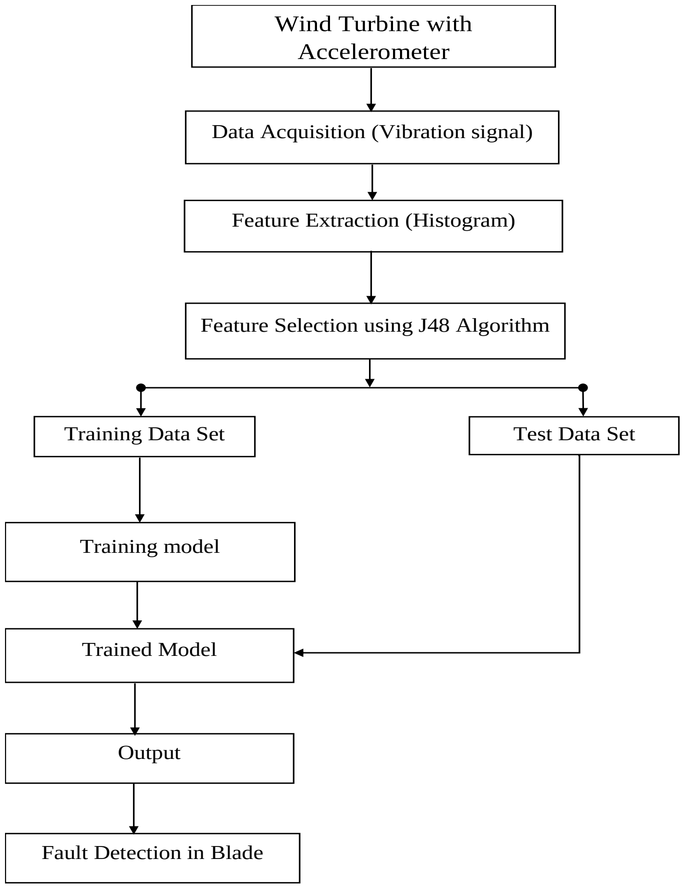


Fig. 1. Methodology.

2. Experimental Studies

The main aim of this study is to classify whether the blades are in good condition or in a defective state. If it is defective, then the objective is to identify the type of fault. The experimental setup and experimental procedure are described in the following subsections [20].

2.1. Experimental setup

The experiment was carried out on a 50W, 12V variable speed wind turbine (MX-POWER, model: FP-50W-12V). The technical parameters of a wind turbine are given in Table 1. The wind turbine was mounted on a fixed steel stand in front of the open circuit wind tunnel outlet. The wind tunnel speed ranges from 5m/s to 15 m/s and acts as a wind source to start the wind turbine. The wind speed was varied continuously in order to simulate the environmental wind condition. The experimental setup is shown in Fig. 2.

Table 1. Technical parameters of wind turbine.

Model	FP-50W-12V
Rated Power	50 W
Rated Voltage	12 V
Maximum Current	4 A
Rated Rotating Rate	850 rpm
Start-up Wind Speed	2.5 m/s
Cut-in Wind Speed	3.5 m/s
Cut-out Wind Speed	15 m/s
Security Wind Speed	40 m/s
Rated Wind Speed	12.5 m/s
Engine	Three-phase permanent magnet generator
Rotor Diameter	1050 mm
Blade Material	Carbon fiber reinforced plastics

Piezoelectric type accelerometer was used as a transducer for acquiring vibration signals. It has high sensitivity for detecting faults. Hence, accelerometers are widely used in condition monitoring. In this case, a uniaxial accelerometer of 500g range, 100 mV/g sensitivity, and resonant frequency around 40 Hz was used. The piezoelectric accelerometer (DYTRAN 3055B1) was mounted on the nacelle near to the wind turbine hub to record the vibration signals using an adhesive mounting technique. It was connected to the DAQ system through a cable. The data acquisition system (DAQ) used was NI USB 4432 model. The DAQ card has five analog input channels with a sampling rate of 102.4-kilo samples per second with 24-bit resolution. The accelerometer is coupled to a signal conditioning unit which consists of an inbuilt charge amplifier and an analogue-to-digital converter (ADC). From the ADC, the vibration signal was taken. These vibration signals were used to extract features through feature extraction technique. One end of the cable is plugged to the accelerometer and the other end to the AIO port of DAQ system. NI - LabVIEW was used to interface the transducer signal and the system (PC).

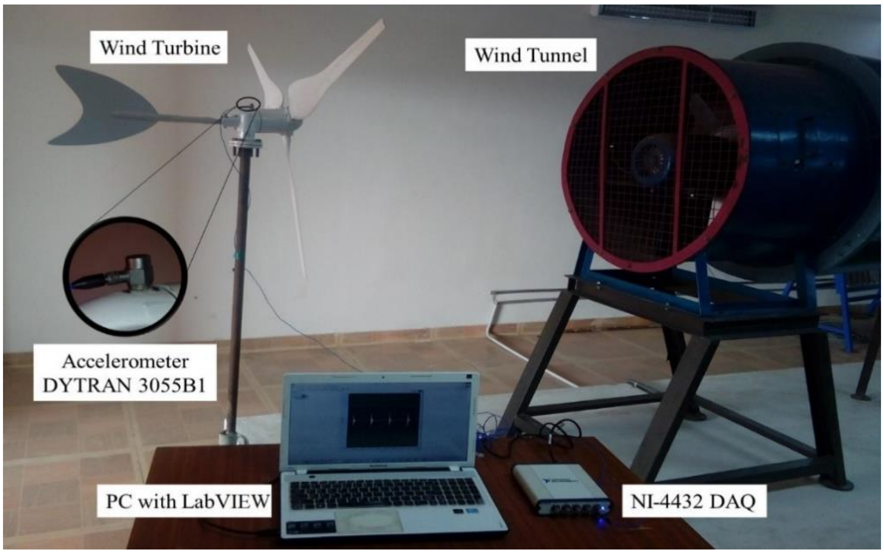


Fig. 2. Wind turbine setup.

2.2. Experimental procedures

In the present study, three-blade variable horizontal axis wind turbine (HAWT) was used. Initially, the wind turbine was considered to be in good condition (free from defects, new setup) and the signals were recorded using an accelerometer. These signals were recorded with the following specifications:

a) **Sample length:** The sample length was chosen long enough to ensure data consistency; and also the following points were considered. Histogram measures are more meaningful when the number of samples is sufficiently large. On the other hand, as the number of samples increases the computation time increases. To strike a balance, a sample length of 10000 was chosen.

b) **Sampling Frequency:** The sampling frequency should be at least twice the highest frequency contained in the signal as per Nyquist sampling theorem. By using this theorem sampling frequency was calculated as 12 kHz (12000 Hz).

c) **Number of samples:** Minimum of 100 (hundred) samples were taken for each condition of the wind turbine blade and the vibration signals were stored in data files [21].

The following faults were simulated one at a time while all other components remain in good condition and the corresponding vibration signals were acquired. Figure 3 shows the different blade fault conditions which are simulated on the blade.

a) **Blade bend (BB):** This fault occurs due to the high-speed wind and complex forces caused by the wind. The blade was made to flap wise bend with 10° angle.

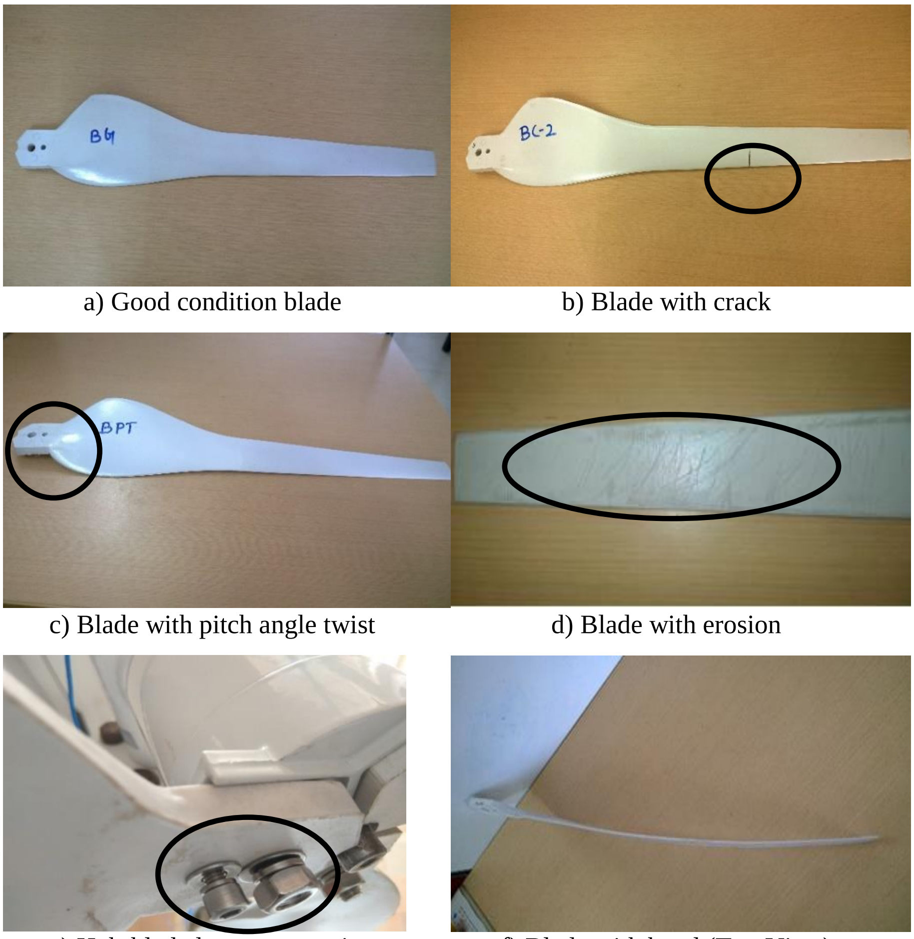
b) **Blade crack (BC-2):** This **occurs** due to foreign object damage on the blade while it is in operating condition. On the blade, 15 mm crack was made.

c) **Blade erosion (BE):** This fault is due to the erosion of the top layer of the blade by the high-speed wind. The smooth surface of the blade was eroded using **emery** sheet (320Cw) to provide an erosion effect on the blade.

d) **Hub-blade loose contact:** This fault generally occurs on a wind turbine blade due to an excessive runtime or usage time. The bolt connecting the hub and blade was made loose to obtain this fault.

e) **Blade pitch angle twist (PAT):** This fault occurs due to the stress on the blade caused by high-speed wind. This makes the pitch get twisted, creating a heavy vibration to the framework. To attain this fault, blade pitch was twisted about 12° angle with respect to the normal blade condition.

From Fig. 4., the vibration signals (sample number vs amplitude) are shown which were taken for different conditions of the wind turbine blade (good condition blade, blade bend, blade erosion, hub-blade loose connection, blade crack and pitch angle twist).

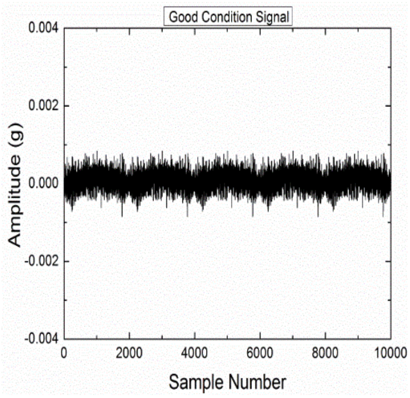


a) Good condition blade b) Blade with crack

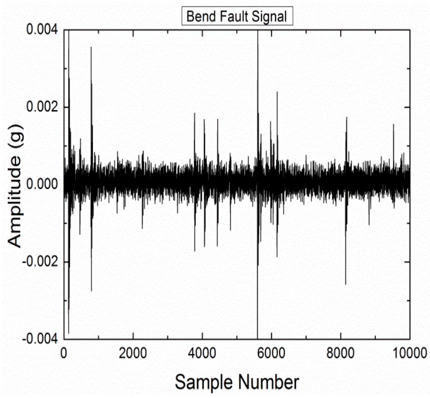
c) Blade with pitch angle twist d) Blade with erosion

e) Hub-blade loose connection f) Blade with bend (Top View)

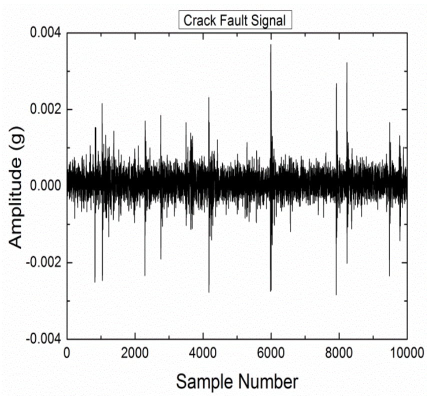
Fig. 3. Various blade fault conditions.



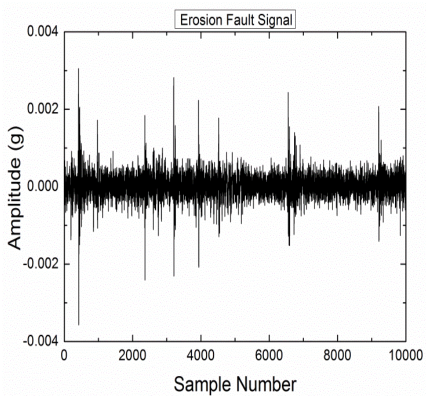
a) Good condition signal plot



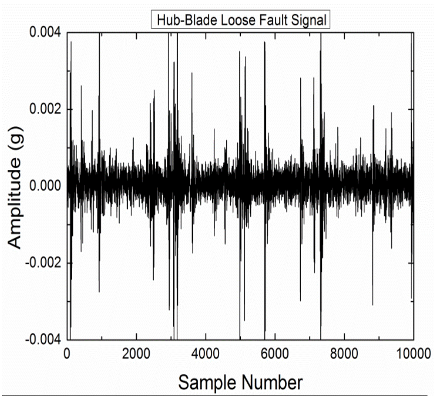
b) Bend fault condition signal plot



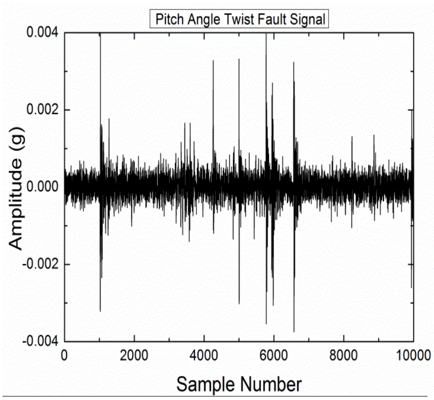
c) Crack fault condition signal plot



d) Erosion fault condition signal plot



e) Hub-blade loose fault condition signal plot



f) Pitch angle twist fault condition signal plot

Fig. 4. Condition signal plot.

3. Feature Extraction

The vibration signals were obtained for good and other faulty conditions of the blades. If the time domain sampled signals are given directly as inputs to a classifier, then the number of samples should be constant. The number of samples obtained is the function of rotation of the blade speed. Hence, it cannot be used directly as the input to the classifier. However, a few features must be extracted before the classification process. The histogram was used as a feature extracting tool in this study. The reason behind choosing the histogram method for feature extraction is because it allows the viewers to easily compare the data and also they work well with large ranges of information or samples. They also provide a more actual form of consistency, as the intervals are always equal, a factor that allows easy data transfer from frequency tables to histograms. Hence, the histogram is preferred for feature extraction.

Feature extraction involves reducing a number of resources required to describe a large set of data. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power, also it may cause a classification algorithm to over-fit to training samples and generalize poorly to new samples. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy. From the noted vibration signals, the needed feature is taken and that features are denoted as histogram features. There are two main factors to be considered in the selection of bins they are, bin range and bin width [22].

Bin is the subrange used for grouping the data. Suppose, we are interested in the distribution of the marks of the students in a class then we have sub ranged like 0-10, 11-20, 21-30…91-100. Each subrange can be called a bin. To construct a histogram, the first step is to "bin" the range of values, that is, divide the entire range of values into a series of intervals and then count how many values fall into each interval. The bins are usually specified as consecutive, non-overlapping intervals of a variable. The bins (intervals) must be adjacent and are often (but are not required to be) of equal size.The bin range must be from lowest of minimum amplitude (-0.017988) to the extreme of maximum amplitude (0.024833) of all the six classes (good, bend, crack, erosion, loose and PAT). The number of bins for the fault diagnosis of wind turbine blade has been attained by carrying out a sequence of trials using a J48 algorithm with a different number of bins. Initially, the range of bin is separated into two equivalent portions. That is to say, the number of bins utilized is two. The two histogram features, to be specific, *X1* and *X2* are extracted and the relating classification accuracy is additionally acquired by using the J48 algorithm. The approach and methodology of performing the same using J48 algorithm are clarified in Section 4. A set of related trails is done with various numbers of bins from 2, 3, 4, 5 to 100 and the corresponding results are shown in Fig. 5.

From Fig. 5, bin size 77 has been chosen since the classification accuracy of bin 77 was found to be 92%. A set of 77 starting from *X1, X2… X77* were extracted from the vibration signals and these are denoted as histogram features. The amplitude ranges from -0.017988 to 0.024833. For further study, rather than utilizing vibration signals directly, the histogram features extracted from vibration signals are utilized. The procedure of calculating applicable parameters of the signals that represent the data contained in the signal is called feature extraction. Histogram analysis of vibration signals yields distinctive parameters. All the extracted histogram features, *X1* to *X77* extracted from the vibration signals may not contain the needed information for classification. The applicable ones are selected using the J48 algorithm.

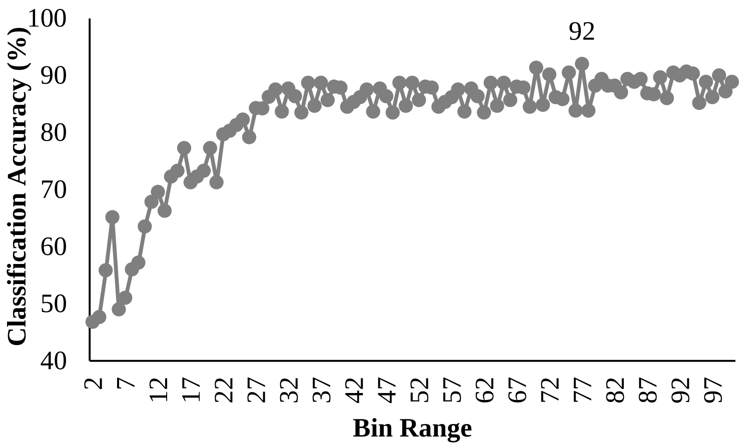


Fig. 5. Bin range vs. classification accuracy.

4. Feature Selection

Feature selection is the process of selecting a subset of relevant features (variables, predictors) for use in model construction. Feature selection techniques are used for four reasons:

- Simplification of models to make them easier to interpret by researchers/users
- Shorter training times
- To avoid the curse of dimensionality
- Enhanced generalization by reducing overfitting (formally, reduction of variance)

The central premise when using a feature selection technique is that the data contains many features that are either redundant or irrelevant and can thus be removed without incurring much loss of information. Redundant or irrelevant features are two distinct notions, since one relevant feature may be redundant in the presence of another relevant feature with which it is strongly correlated. From the extracted features (77), the most contributing features are selected using feature selection process. For feature selection, J48 decision tree algorithm is used. J48 decision tree algorithm is adapted from the C4.5 algorithm in WEKA [23]. It consists of a number of branches, one root, a number of nodes, and a number of leaves. One branch is a chain of nodes from the root to a leaf, and each node involves one attribute. The occurrence of an attribute in a tree provides information about the importance of the associated attribute. A decision tree is a tree based knowledge representation methodology used to represent classification rules. J48 decision tree algorithm is a widely used one to construct decision trees.

The procedure of forming the decision tree and exploiting the same for feature selection is characterized by the following:

a) The set of features available at hand forms the input to the algorithm; the output is the decision tree.

b) The decision tree has leaf nodes, which represent class labels, and other nodes associated with the classes being classified.

c) The branches of the tree represent each possible value of the feature node from which they originate.

d) The decision tree can be used to classify feature vectors by starting at the root of the tree and moving through it right through to a leaf node, which provides a classification of the instance, is identified.

e) At each decision node in the decision tree, one can select the most useful feature for classification using appropriate estimation criteria. The criterion used to identify the best feature invokes the concepts of entropy reduction and information gain.

Information gain measures how well a given attribute separates the training examples according to their target classification. The measure is used to select the candidate among the features at each step while growing the tree. Information gain is the expected reduction in entropy caused by portioning the samples according to this feature.

Information gain (S, A) of a feature A relative to a collection of examples S, is defined as:

$$Gain\ (S,A) = Entropy(S) - \sum_{v \in Value(A)} \times \frac{|S_v|}{|S|} Entropy(S_v) \tag{1}$$

where *Value (A)* is the set of all possible values for attribute *A*, and S_v is the subset of *S* for which feature *A* has value *v*. Note the first term in the equation for gain is just the entropy of the original collection *S* and the second term is the expected value of the entropy after *S* is partitioned using feature *A*. The expected entropy described by the second term is simply the sum of the entropies of each subset S_v, weighed by the fraction of samples $|S_v|/|S|$ that belong to S_v. *Gain (S, A*) is the expected reduction in entropy caused by knowing the value of feature *A*. Entropy is a measure of homogeneity of the set of examples and it is given by

$$Entropy(S) = \sum_{i-1}^{c} -P_i \log_2 P_i \tag{2}$$

where *c* is the number of classes, P_i is the proportion of *S* belonging to class 'i'.

The J48 decision tree algorithm has been applied to the problem for feature selection process. The input to the algorithm is the set of histogram features described above and output of the decision tree shown in Fig. 6. It clearly shows that the top node is the best node for classification. The other features in the nodes of a decision tree are seen in descending order of significance. It is to be mentioned here that only features that contribute to the classification appear in the decision tree. The features which have less of a discriminating capability can be consciously discarded by deciding on the threshold. This concept is made clear for selecting relevant features. The algorithm identifies the relevant features for the purpose of classification from the given training data set, and thus reduces the domain knowledge required to select good features for pattern classification problem. Referring from Fig. 6, one can identify the most dominating features to represent the blade conditions are *X33, X34, X35, X32,* and *X36*.

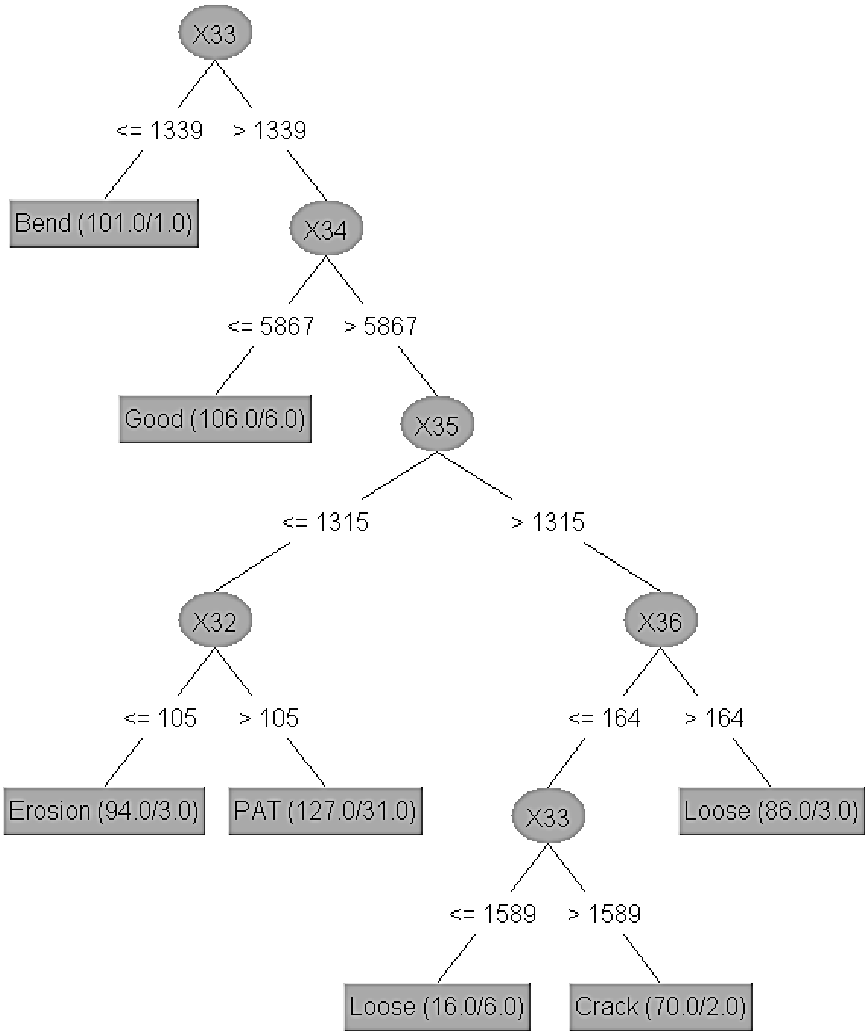


Fig. 6. J48 Tree classification for feature selection.

5. Feature Classification

After the feature selection, the fault classification was carried out using, sequential minimal optimization (SMO) algorithm, the simple logistic algorithm (SLA), multilayer perceptron (MLP), the logistic algorithm (LA) and radial basis function (RBF). These algorithms are explained in this section below.

5.1. Sequential minimal optimization (SMO)

Sequential minimal optimization (SMO) is a classifier for undertaking the quadratic programming issue that emerges in the training of support vector machines. Consider a binary classification issue with a data set $(x_1, y_1), ..., (x_n, y_n)$, where x_i is an input vector and $y_i \in \{-1, +1\}$ is a binary name associating with it. Sequential minimal optimization is an iterative calculation for taking care of the optimization issue [24]. Sequential minimal optimization breaks this issue into a progression of least conceivable sub-issues, which are then elucidated logically. Due to the linear equality requirement, including the Lagrange multipliers α_i, the least conceivable issue includes two such multipliers. At that point, for any two multipliers α_1 and α_2, the requirements are narrowed to:

$$0 \leq \alpha_1, \alpha_2 \leq C \tag{3}$$

$$y_1\alpha_1 + y_2\alpha_2 = k \tag{4}$$

where, k is the negative of the sum over. The main advantage of SMO classifier is that it will predict the direction of the fault improvement and it helps the formulator to decide optimum conditions for the formulation and process.

5.2. Simple logistic algorithm (SLA)

The simple logistic classifier is comparable to linear regression, with the exception of that the dependent variable is nominal, not an estimation. One objective is to see whether the likelihood of getting a specific estimate of the nominal variable is connected with the estimation variable [25]. The other objective is to forecast the likelihood of getting a specific estimate of the nominal variable, given the estimation variable. The simple logistic regression finds the mathematical equation that best forecasts the estimation of the Y variable for every estimation of the X variable. The logistic regression is not relatively the same as linear regression is that one does not quantify the Y variable specifically; it is rather the likelihood of acquiring a specific estimation of a nominal variable. The equation is

$$ln\left[\frac{Y}{1-Y}\right] = a + bX \tag{5}$$

where a is the slope and b are the intercepts of the best-fitting mathematical equation in a logistic regression by means of the maximum-likelihood technique, relatively than the least-squares technique. The main advantage of SLA is that it can provide the independent observation of the data set and it has fast computation.

5.3. Multilayer perceptron (MLP)

A multilayer perceptron (MLP) is a feedforward artificial neural network model that plots sets of data onto an arrangement of suitable yields. A multilayer perceptron contains different layers of hubs in an engaged outline, with individual layer totally connected to the following one. Aside from the input hubs, the individual hub is a neuron or preparing component with a nonlinear initiation capability. Multilayer perceptron utilizes a directed learning technique called backpropagation for instructing the system. A multilayer perceptron is a change of the standard linear perceptron and can separate information that is not linearly separable. The basic concept of a single perceptron was introduced by Rosenblatt in 1958 [26]. The perceptron computes a single output from multiple real-valued inputs by forming a linear combination according to its input weights and then possibly putting the output through some nonlinear activation function. Mathematically, this can be written as

$$y = \varphi\left(\sum_{i=1}^{n} \omega_i x_i + b\right) = \varphi(\omega^T x + b) \tag{6}$$

where ω denotes the vector of weights, X is the vector of inputs, b is the bias and φ are the activation function. The main advantage of MLP is that it can train the model by back propagation algorithm to perform any mapping between the input and the output signal.

5.4. Logistic algorithm

In this study, logistic regression is used for classification in this multiclass problem. The logistic algorithm used to measure the connection between the definite variable and one or more free variables by assessing probabilities utilizing a logistic capacity, which is the combined logistic distribution [27]. In this way, it treats the same arrangement of issues as the probit function utilizing equivalent strategies, with the last utilizing an overall normal distribution curve as an alternative. Equally, in the dormant variable elucidations of these two strategies, the primary expects a standard logistic distribution of errors and the second a standard typical normal distribution of errors. The logistic algorithm is represented by

$$\log \frac{p\,(x)}{1-p(x)} = \beta_0 + x \cdot \beta \tag{7}$$

where $p\ (x)$ is the linear function of x, the dependent variable is x and the independent variable is β. The main advantage of LA is that it can able to predict the outcome of the categorical dependent variable and has fast computation.

5.5. Radial basis function (RBF) network

It is a real-valued function whose value depends only on the distance from the origin, so that $\phi\ (x) = \phi\ (||x||)$; or alternatively on the distance from any other point c, called a centre, so that $\phi\ (x,c) = \phi\ (||x\text{-}c||)$. It can also be interpreted as a relative simple single-layer type of artificial neural network called a radial basis function network, with the radial basis functions taking on the role of the activation functions of the network.

$$y(X) = \sum_{i=1}^{N} \omega_i \emptyset\, (\|X - X_i\|) \tag{8}$$

It can be exhibited that any consistent function on a minimal disruption can a standard to be added with discretionary precision by an aggregate of this structure if a sufficiently large number of radial basis functions is used The approximant $y(\mathbf{x})$ is differentiable with respect to the weights w_i. The weights could thus be learned using any of the standard iterative methods for neural networks [28]. Utilizing radial basis functions as a part of this way produces a sensible interjection approach gave that the fitting set has been picked such that it covers the whole range deliberately where equidistant information points are ideal. Though, without a polynomial term that is orthogonal to the radial basis functions, measures outside the fitting set have a tendency to perform ineffectively. The main advantage of RBF is that it uses a small number of locally tuned units and is adaptive in nature.

6. Results and Discussion

The vibration signals were noted for good condition and faulty blade conditions using DAQ. The total number of signals collected is 600; out of which 100 samples were from good condition blade. For different faults such as blade bend, erosion, blade crack, hub-blade loose connection, pitch angle twist, 100 samples from every condition were noted. J48 decision tree algorithm was used to select the best contributing histogram features from bin size 77. From Figs. 5 and 6, the selected features are given as the input to the classifier to determine the classification accuracy with respect to faults created by the wind turbine blade. From Fig. 7, the simple logistic algorithm (SLA) gives the maximum classification accuracy of 93.67% when compared to other classifiers. In the simple logistic algorithm, the complexity parameter (C) was fixed to be 1. The filter type is made as normalize training data and the kernel is chosen as a polynomial. The confusion matrix of the simple logistic algorithm (SLA) is shown in Table 2. In confusion matrix, the diagonal element represents the correctly classified instance and the others are misclassified [29].

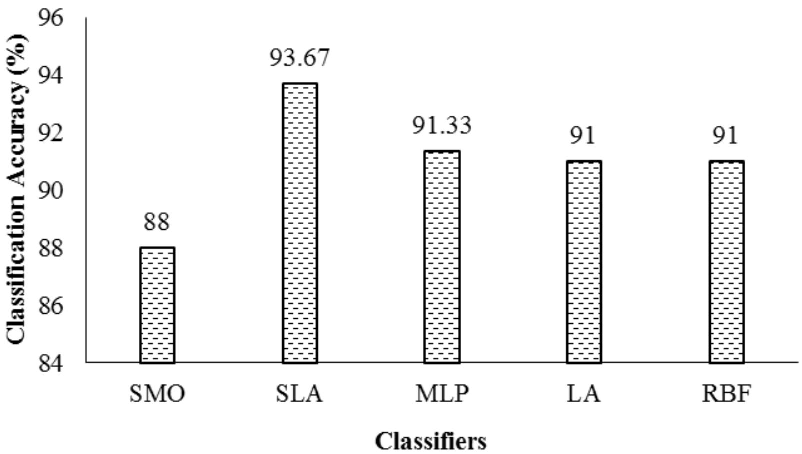


Fig. 7. Overall classification accuracy of the classifiers.

From simple logistic algorithm, the kappa statistics were found to be 0.924. It is used to measure the arrangement of likelihood with the true class. The mean absolute error was found to be 0.0356. It is a measure used to measure how close forecasts or prediction are with the ultimate result. The root mean square error was found to be 0.1263. It is a quadratic scoring rule which processes the average size of the error. The detailed class wise accuracy is shown in Table 3. Of 600 samples, 562 samples are correctly classified (93.67%) and remaining 38 are misclassified (6.33%). The time taken to build the model is about 1.05 seconds; hence, this can use in real time for the fault detection on the wind turbine blade. The model was tested in 10-fold cross validation. Cross-validation is a technique to evaluate predictive models by partitioning the original sample into a training set to train the model, and a test set to evaluate it.

In 10-fold cross-validation, the original sample is randomly partitioned into 10 equal size subsamples. Of the 10 subsamples, a single subsample is retained as the validation data for testing the model, and the remaining 9 subsamples are used as training data. The cross-validation process is then repeated 10 times (the folds), with each of the 10 subsamples used exactly once as the validation data. The 10 results from the folds can then be averaged (or otherwise combined) to produce a single estimation. The advantage of this method is that all observations are used for both training and validation, and each observation is used for validation exactly once [30].

Table 2. Confusion matrix for simple logistic algorithm (SLA).

Blade conditions	Good	Bend	Crack	Erosion	Loose	PAT
Good	**95**	0	1	0	4	0
Bend	2	**97**	0	1	0	0
Crack	1	0	**88**	2	6	3
Erosion	0	0	5	**95**	0	0
Loose	5	0	1	0	**94**	0
PAT	0	0	6	1	0	**93**

Table 3. Class wise accuracy of simple logistic algorithm (SLA).

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC area
Good	0.95	0.016	0.922	0.95	0.936	0.996
Bend	**0.97**	**0.000**	**1.000**	**0.97**	**0.985**	**0.983**
Crack	0.88	0.026	0.871	0.88	0.876	0.989
Erosion	**0.95**	**0.008**	**0.96**	**0.95**	**0.955**	**0.999**
Loose	0.94	0.02	0.904	0.94	0.922	0.996
PAT	**0.93**	**0.006**	**0.969**	**0.93**	**0.949**	**0.997**

From class wise accuracy in Table 3, the properties like true positive rate (TP), false positive rate (FP), precision, recall, F-Measure, receiver operating characteristics (ROC) area are determined [31]. TP is also called as sensitivity which used to predict the ratio of positives which are correctly classified as faults. FP is commonly described as a false alarm in which the result that shows a given fault condition has been achieved when it really has not been achieved. The true positive (TP) rate should be close to 1 and the false positive (FP) rate should be close to 0 to propose the classifier is a better classifier for the problem classification. In the simple logistic algorithm (SLA), it shows that the TP near to 1 and FP close to 0, then one can predict that the classifier we build for the particular problem is very much effective for the fault diagnosis problem.

Precision is the ratio of correctly classified instances for those instances that have been classified as positive. The recall is merely equal to sensitivity in which the information retrieval is the fraction of the faults that are relevant to the query that are successfully retrieved. F-measure is defined as the equivalent resistance formed by sensitivity and precision positioned in parallel phase. ROC is a graphical representation that demonstrates the performance of a classifier as its discrimination threshold is varied. The classifier error chart is shown in Fig. 8. Here the squared dots represent the misclassification and the '*x*' denotes the correct classification.

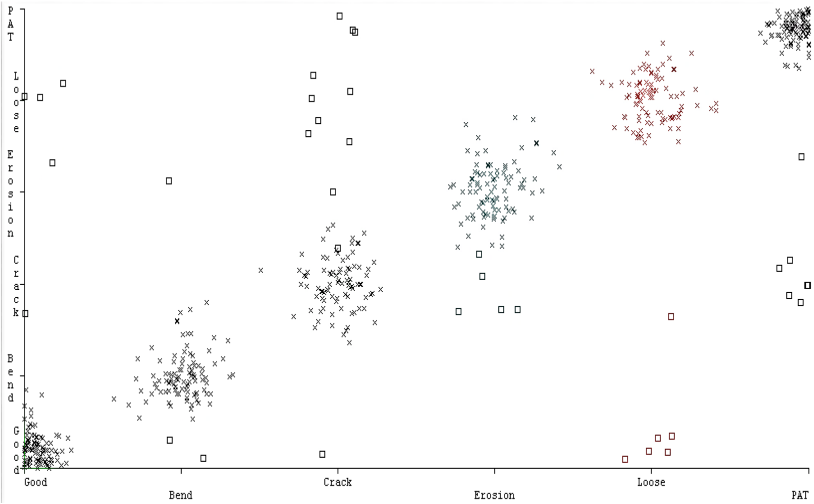


Fig. 8. Classifier errors (classification vs. misclassification).

7. Conclusion

The wind turbine is a very important structure in extracting wind energy from the accessible wind. This paper displayed an algorithm based classification of vibration signals for the evaluation of the wind turbine blade conditions. From the acquired vibration data, five models were developed using data modeling techniques. The model was tested in 10-fold cross validation. All the classifiers were compared with respect to their types and maximum correctly classified instances were found to be 93.67% for the simple logistic algorithm (SLA). The error rate is relatively less and may be considered for the blade fault diagnosis. Hence, the simple logistic algorithm (SLA) can be practically used for the condition monitoring of wind turbine blade to reduce the downtime and to maximize the usage of wind energy. The methodology and algorithm suggested in this paper can be potentially used for any kind of wind turbine blade to diagnose the blade fault with minimal modification.

References

1. Liu, W.Y.; Tang, B.P.; Han. J.G.; Lu. X.N.; Hu. N.N.; and He. Z.Z. (2015). The structure healthy condition monitoring and fault diagnosis methods in wind turbines: A review. *Renewable and Sustainable Energy Reviews*, 44, 466-472.

2. Chehouri, A.; Younes, R.; Ilinca, A.; and Perron, J. (2015). Review of performance optimization techniques applied to wind turbines. *Applied Energy*, 142, 361-388.
3. Ciang, C.C.; Lee. J.R.; Bang. H.J. (2008). Structural health monitoring for a wind turbine system: A review of damage detection methods. *Measurement Science and Technology*, 19(12), 1-20.
4. Hameed, Z.; Hong, Y.S.; Cho, Y.M.; Ahn, S.H.; and Song. C.K. (2009). Condition monitoring and fault detection of wind turbines and related algorithms: A review. *Renewable and Sustainable energy reviews*, 13(1), 1-39.
5. Tummala, A.; Velamati, R.K.; Sinha, D.K.; Indraja, V.; and Krishna, V.H. (2016). A review on small scale wind turbines. *Renewable and Sustainable Energy Reviews*, 56, 1351-1371.
6. Joshuva, A.; and Sugumaran, V. (2016). Fault diagnostic methods for wind turbine: A review. *ARPN Journal of Engineering and Applied Sciences*, 11(7), 4654-4668.
7. Amirat, Y.; Benbouzid, M.E.; Al-Ahmar, E.; Bensaker, B.; and Turri, S. (2009). A brief status on condition monitoring and fault diagnosis in wind energy conversion systems. *Renewable and Sustainable Energy Reviews*, 13(9), 2629-2636.
8. Frost, S.A.; Goebel, K.; and Obrecht. L. (2013). Integrating structural health management with contingency control for wind turbines. *IJPHM Special Issue* on *Wind Turbine PHM* (Color). 11.
9. Ataya, S.; and Ahmed, M.M. (2013). Damages of wind turbine blade trailing edge: Forms, location, and root causes. *Engineering Failure Analysis*, 35, 480-488.
10. Dervilis, N.; Choi, M.; Taylor, S.G.; Barthorpe, R.J.; Park. G.; Farrar, C.R.; and Worden, K. (2014). On damage diagnosis for a wind turbine blade using pattern recognition. *Journal of Sound and Vibration*, 333(6), 1833-1850.
11. Roth-Johnson, P.; Wirz, R.E.; and Lin, E. (2014). Structural design of spars for 100-m biplane wind turbine blades. *Renewable Energy*, 71, 133-155.
12. Lee, J.-K.; Park, J.-Y.; Oh, K.-Y.; Ju, S.-H.; and Lee, J.-S. (2015). Transformation algorithm of wind turbine blade moment signals for blade condition monitoring. *Renewable Energy*, 79, 209-218.
13. Saravanakumar, R.; and Jena, D. (2015). Validation of an integral sliding mode control for optimal control of a three blade variable speed variable pitch wind turbine. *International Journal of Electrical Power & Energy Systems*, 69, 421-429.
14. Vučina, D.; Marinić-Kragić, I.; and Milas, Z. (2016). Numerical models for robust shape optimization of wind turbine blades. *Renewable Energy*, 87, 849-862.
15. Pourrajabian, A.; Afshar, P.A.; Ahmadizadeh. M.; and Wood, D. (2016). Aero-structural design and optimization of a small wind turbine blade. *Renewable Energy*, 87(Part 2), 837-848.
16. de Bessa, I.V.; Palhares, R.M.; D'Angelo, M.F.; and Chaves Filho, J.E. (2016). Data-driven fault detection and isolation scheme for a wind turbine benchmark. *Renewable Energy*, 87(Part 1), 634-645.

17. Rafiee, R.; Tahani, M.; and Moradi, M. (2016). Simulation of aeroelastic behavior in a composite wind turbine blade. *Journal of Wind Engineering and Industrial Aerodynamics*, 151, 60-69.
18. Hoell, S.; and Omenzetter, P. (2016). Optimal selection of autoregressive model coefficients for early damage detectability with an application to wind turbine blades. *Mechanical Systems and Signal Processing*, 70, 557-577.
19. Rezaei, M.M.; Behzad, M.; Moradi, H.; and Haddadpour, H. (2016). Modal-based damage identification for the nonlinear model of modern wind turbine blade. *Renewable Energy*, 94, 391-409.
20. Joshuva, A.; and Sugumaran, V. (2017). A data driven approach for condition monitoring of wind turbine blade using vibration signals through best-first tree algorithm and functional trees algorithm: A comparative study. *ISA Transactions*, 67, 160-172.
21. Joshuva, A.; and Sugumaran, V. (2017). Classification of various wind turbine blade faults through vibration signals using hyperpipes and voting feature intervals algorithm. *International Journal of Performability Engineering*, 13(3), 247-258.
22. Joshuva, A.; Sugumaran, V.; and Amarnath, M. (2015). Selecting kernel function of Support Vector Machine for fault diagnosis of roller bearings using sound signals through histogram features. *International Journal of Applied Engineering Research*, 10(68), 482-487.
23. Joshuva, A.; and Sugumaran, V. (2016). Wind turbine blade fault diagnosis using vibration signals through decision tree algorithm. *Indian Journal of Science and Technology*, *9*(48).
24. Keerthi, S.S.; Shevade, S.K.; Bhattacharyya, C. and Murthy, K.R. (2001). Improvements to Platt's SMO algorithm for SVM classifier design. *Neural Computation*, 13(3), 637-649.
25. Ren, X.; and Malik, J. (2003). Learning a classification model for segmentation. *In Computer Vision Proceedings, Ninth IEEE International Conference*, 10-17.
26. Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6), 386.
27. Hilbe, J.M. (2009). Logistic regression models. *CRC Press*.
28. Yingwei, L.; Sundararajan, N.; and Saratchandran, P. (1998). Performance evaluation of a sequential minimal radial basis function (RBF) neural network learning algorithm. *Neural Networks*, 9(2), 308-318.
29. Joshuva, A.; and Sugumaran, V. (2017). A comparative study of Bayes classifiers for blade fault diagnosis in wind turbines through vibration signals. *Structural Durability and Health Monitoring* (*SDHM*), 12(1), 69-90
30. Joshuva, A.; and Sugumaran, V. (2015). Speech recognition for humanoid robot. *International Journal of Applied Engineering Research*, 10(68), 57-60.
31. Joshuva, A.; Sugumaran; V., Amarnath; M. and Lee, S. K. (2016). Remaining life-time assessment of gear box using regression model. *Indian Journal of Science and Technology*, 9(47).