Journal Paper

**“Dirt and Mud Detection and Diagnosis on a Wind Turbine Blade
employing Guided Waves and Supervised Learning Classifiers”**

- Reliability Engineering & System Safety 2018-

Alfredo Arcos Jimenez
Ingenium Research Group, Universidad de Castilla-La Mancha
Alfredo.Arcos@alu.uclm.es

Carlos Quiterio Gómez Muñoz
European University of Madrid, Spain.
carlosquiterio.gomez@universidadeuropea.es

Fausto Pedro García Márquez
Ingenium Research Group, Universidad de Castilla-La Mancha
FaustoPedro.Garcia@uclm.es

Cite as: Jiménez, A. A., Muñoz, C. Q. G., & Márquez, F. P. G. (2018). Dirt and mud detection and diagnosis on a wind turbine blade employing guided waves and supervised learning classifiers. Reliability Engineering & System Safety.

Dirt and Mud Detection and Diagnosis on a Wind Turbine Blade employing Guided Waves and Supervised Learning Classifiers

Alfredo Arcos Jiménez¹, Carlos Quiterio Gómez Muñoz^{2,*}, Fausto Pedro García Márquez¹

¹ Ingenium Research Group, Castilla-La Mancha University, Spain.

alfredo.arcos@alu.uclm.es; faustopedro.garcia@uclm.es

² European University of Madrid, Spain.

carlosquiterio.gomez@universidadeuropea.es

Abstract.

Dirt and mud on wind turbine blades (WTB) reduce productivity and can generate stops and downtimes. A condition monitoring system based on non-destructive tests by ultrasonic waves was used to analyse it. This paper employs an approach that considers advanced signal processing and machine learning to determine the thickness of the dirt and mud in a WTB. Firstly, the signal is filtered by Wavelet transform. FE and Feature selection (FS) are employed to remove non-useful data and redundant features. FS selects the number of the most significant terms of the model for fault detection and identification, reducing the dimension of the dataset. Pattern recognition is carried out by the following supervised learning classifiers based on statistical models to calculate and classify the signal depending on the fault: Ensemble Subspace Discriminant; k-Nearest Neighbours; Linear Support Vector Machine; Linear Discriminant Analysis; Decision Trees. Receiver Operating Characteristic analysis is used to evaluate the classifiers. Neighbourhood Component Analysis has been employed in feature selection. Several case studies of mud on the WTB surface have been considered to test and validate the approach. Autoregressive (AR) model and Principal Component Analysis (PCA) have been employed to FE. The results provided by PCA show an improvement on the AR results. The novelty of this work is focused on applying this approach to detect and diagnose mud and dirt in WTB.

Keywords: Macro Fiber Composite, Wavelet Transforms, Machine Learning, Guided Waves, Wind Turbine Blade.

1 Introduction

The wind energy industry is producing more power according to the Kyoto Protocol [1-4]. Wind turbines are subjected to high mechanical loads, and require an optimal maintenance policy to provide a cost-effective power output and minimise life cycle costs [5,6].

Wind turbine blade (WTB) is a critical component in an aerogenerator. The WTB movement is generated by the pressure difference that produces an air foil between the upper and lower sides of the WTB [7]. WTBs are designed to generate a low pressure and suction zone on the upper side of the WTB, and a high-pressure zone on the underside of the WTB [8]. The soiling generates turbulence that reduces the efficiency of the WTB, i.e. less production and profit for this industry.

Soiling on wind turbine WTBs is present mainly in the edges, caused by insects, mud, dust, salt, ice, etc. Soiling can also be caused by intermittent rain, low humidity, farming activities, flying insects, etc., where the geographical situation and the season also affect the severity of soiling [9].

Figure 1 Accumulation of mud and mosquitoes on a WTB [10].

The soiling on **the leading edge** reduces the WTB's performance, energy production and increases the emitted noise. It also produces disturbances in the air flow layers, close to the surface, generating a turbulent regime in the air flow, favouring the emergence of the 'stall' phenomenon. Kinetic energy is decreased and, therefore, mechanical rotation energy and the electricity generated [11]. In addition, it produces false alarms due to the reduction in the performance of the wind turbine. The cleaning task of a WTB is an expensive operation that requires special equipment such as cranes, qualified staff, etc., which are necessary to stop wind turbines [12,13].

The condition-based maintenance (CBM) is an advanced maintenance strategy based on monitoring the condition of a component. It is possible to obtain measurements of condition monitoring of the WTB, including data from vibration, acoustic emission, infrared images, etc. The CBM objective is to optimize maintenance activities and reduce costs [14-19].

Non-destructive testing (NDT) is used in CBM [20-22] and failure prognostics [23]. Ultrasonic guided waves is a NDT technique employed for the inspection of WTBs [24,25]. Lamb waves are ultrasonic waves that are utilised in structures as plates or shell shape, being sensitive to changes in the material, e.g. soiling. The longitudinal modes of Lamb waves produce a particle displacement which is perpendicular to the direction of wave propagation. Longitudinal Lamb waves are composed of two different vibration modes: symmetric and anti-symmetric modes (Figure 2).

Figure 2 Propagation modes of Lamb waves: Symmetric (left) and Anti-symmetric (right) [26].

The mud and dirt accumulated on a WTB's surface execute a force on the surface particles that opposes the free movement of these particles, resulting in small variations in the ultrasonic signals. The use of advanced signal processing and pattern recognition methods for detection and diagnosis is required for CBM [27].

Farrar and Sohn applied pattern recognition in Structure Health Monitoring (SHM), where the approach is summarised in the four stages shown in Figure 3 [28].

Figure 3. Methodology SHM.

The approach discussed in this paper filters the signals by Wavelet transform [29]. Dimensionality reduction techniques have been developed to remove noise and redundant features by Feature Extraction (FE) and Feature Selection (FS). FE contains the characteristic information of the signal that distinguishes between a fault or free of fault in a structure [30]. FS selects the number of terms that are more appropriate for the model created for fault detection and diagnosis, and the dimension of the dataset is reduced [28,31]. Neighbourhood Component Analysis (NCA) has been employed in FS. Finally, classifiers, based on statistical models, are employed to calculate and classify the data depending on the fault. These classifiers are grouped into two categories: Supervised Learning (SL) and Unsupervised Learning [32]. The following SL classifiers are used in this paper: Ensemble Subspace Discriminant; k-Nearest Neighbours; Linear Support Vector Machine; Linear Discriminant Analysis; Decision Trees. Autoregressive (AR) model and Principal Component Analysis (PCA) have been employed to Feature Election. They are evaluated and validated by receiver operating characteristic analysis.

The novelty of this work is focused on applying the above approach to detect and diagnose mud and dirt on WTB.

2 State of the art

The detection and diagnosis of dust and mud on wind turbine blades has not been sufficiently covered in literature to date. It is being done indirectly through dataset analysis from Supervisory Control and Data Acquisition (SCADA) [33]. Trend analysis techniques for the main parameters of a wind turbine, e.g. power, pitch angle, rotational speed or wind speed, are employed to detect dirt accumulation [34]. Direct detection techniques have been studied using passive sensors, e.g. multispectral, short-wave infrared, mid-wave infrared, long-wave infrared, polarisation or stereo sensors [35]. Images taken from the sensors, through their visible spectral bands, are used to detect mud. Milinda *et al.* propose a new algorithm for detecting types of dirt on textured surfaces using camera images for domestic robotics [36].

Artificial intelligence and Machine Learning (ML) are employed in pattern recognition [37]. Pérez-Ortiz *et al.* present a review of ML algorithms for pattern recognition in renewable energies [38]. Dervilis *et al.* apply pattern recognition for WTB failure diagnosis, considering Principal Component Analysis (PCA), Non-Linear Principal Component Analysis (NLPCA), Artificial Neural Networks (ANN) with Radial Base Functions (RBF) in vibration signals [39]. Pavlopoulou *et al.* consider guided ultrasonic waves with post-processing techniques by FE and FS by PCA and NLPCA in a case study based on the SHM online of scarf repairs in composite laminates [40].

The signal is firstly filtered by wavelet transform in this paper. PCA [41] and AR by Yule-Walker [42] are used for FE. There is a large number of references where this methodology is applied in the field of SHM [43-45]. The main objective is to validate an FE method for the detection and diagnosis of mud in WTB. NCA has been applied to reduce dimensionality and avoid overfitting. NCA is an FS method that provides a good performance and can reduce the number of features with high accuracy.

3 Case study: Experimental procedure

A layer of mud was applied to the surface of a WTB's section to simulate dirt (see Figure 4), composed of a mixture of sand and a solution of water and glue.

Figure 4 Mud on a WTB.

The first thickness of the mud layer was set between 5 and 6.5 mm. Then, the thickness was reduced gradually. Five levels of mud on the WTB were considered, and they are shown in Table 1. The measurements were taken using a high precision digital calliper.

TABLE 1.
THICKNESS OF MUD LAYER.

Level of mud	Layer thickness
1	5-6.5mm (max)
2	3.5-5mm
3	2-2.5mm
4	0.5-1.5mm
5	0 mm

Guided waves were generated using Macro Fiber Composite (MFC) transducers [46] (see Figure 5). The technique used is called 'pitch and catch', where a short ultrasonic pulse is emitted by the MFC transmitter (Tx), and it is collected by the MFC sensor (Rx). The optimal emission pulse for this case study was a six-cycle pulse cycle modulated with a Hanning window. The direct pulse transmitted from the Tx to Rx detects and diagnoses the dirt or mud on the WTB. According to the manufacturer, MFC must be applied to the surface ensuring they make a correct contact with the WTB. It has been demonstrated that an excess of bonding can impede the transmission of movement from MFC to the WTB. The transducers are separated from the edges of the WTB to avoid unwanted echoes from the edges.

Figure 5 Scheme of MFC location for mud detection.

Six different excitation frequencies in Tx were employed for each scenario to set the most sensitive frequencies for detecting and diagnosing mud on the surface. The frequencies of the transmitter were: 18 kHz, 20 kHz, 25 kHz, 30 kHz, 40 kHz and 50 kHz. 50 signals were collected from each frequency and each level of mud, and 3570 samples per experiment, with 4 Msamples/s of sample rate. Figure 6 shows all the frequencies employed with the maximum level of mud (level 1).

Figure 6. Set of signals for level 1 for [18, 20, 25, 30, 40, 50] kHz.

4 Detection and diagnosis approach

The signal is conditioned and de-noised to train the classifiers by wavelet transform [24,25]. The number of inputs and the training patterns will determine the performance and architecture of the classifiers. It is necessary to reduce the number of inputs and to maintain the main information of the signal regarding the condition of the WTB. FE were extracted by employing AR, the Yule-Walker method, and PCA. The number of terms for the FS was done by NCA. Akaike

information criterion (AIC) was also employed, but the results obtained were not considered in this paper because they were worse than the results obtained by NCA.

Different classifiers have been employed to detect and diagnose the mud: Ensemble Subspace Discriminant; k-Nearest Neighbours; Linear Support Vector Machine; Linear Discriminant Analysis, and; Decision Trees. A training cross validation has been employed to avoid overfitting. Figure 7 shows the schema of the approach.

Figure 7. Process flow to detect and diagnose mud and dirt in WTB.

4.1 Signal Pre-processing

4.1.1 Denoising by Wavelet

Wavelet Transform is used to denoise the signals to improve the limitations of resolution and the loss of information presented by the Short-Time Fourier Transform or the Fast Fourier Transform [47]. It uses a variable window size to extract the low frequencies (approximations, A) and the high frequencies (details, D) from the original signal. The process of noise filtering with Wavelet transform eliminates certain high frequency bands maintaining the main information of the signal for pattern recognition.

There are several families of Wavelet Transform, e.g. Haar, Daubechies, Biortogonal, Coiflets, Morlet, Symlet, etc. The Daubechies wavelet family is employed because it is more sensitive to sudden changes in the signal [48,49].

Signal denoising is performed by also employing the multilevel 1-D wavelet analysis, Daubechies family. The threshold for the de-noising is obtained by a wavelet coefficients selection rule using a penalization method provided by Birgé-Massart. In contrast to other digital filters, the Wavelet de-noising filter does not produce an unwanted delay in the signal and removes the noise significantly.

4.1.2 FE methods.

AR Method

The output is a linear combination of the past p outputs plus a white noise input in an AR model of order p . The weights on the p past outputs minimize the mean-square prediction error of the auto-regression, given by equation (1).

$$y(t) = \sum_{i=1}^p \phi_i y(t-i) + \xi(t) \quad (1)$$

where $y(t)$ is the time series to be modelled, ϕ_i are the model coefficients, $\xi(t)$ is white noise, and parameter p is the order of the AR model. An AR model of order p , denoted by $AR(p)$, indicates that the current observation depends on p past observations. The coefficients are set by Yule-Walker AR by equation (2) [50,51]:

$$y(i+1) = \sum_{i=1}^p \phi_1 y(i) + \phi_2 y(t-i) + \dots + \phi_p y(i-p+1) + \xi(i+1) \quad (2)$$

PCA

PCA [41] is a method for mapping multidimensional data into lower dimensions with minimal loss of information.

Being $\mathbf{Y}$ a matrix $n \times m$, where n is the number of observations and m is the number of variables, PCA is an optimal factorization of $\mathbf{Y}^{m \times n}$ into two matrices, $\bar{\mathbf{T}}^{n \times f}$, called scores matrix, and $\bar{\mathbf{P}}^{m \times f}$, named loadings matrix, plus a matrix of residuals $\bar{\mathbf{E}}$, where f is the number of factors ($f < m$). PCA is expressed by equation (3):

$$\mathbf{Y} = \bar{\mathbf{T}}\bar{\mathbf{P}}^T + \bar{\mathbf{E}} \quad (3)$$

PCA is a linear mapping of data from $\mathcal{R}^m$ to $\mathcal{R}^f$, that can also be given by equation (4):

$$\bar{\mathbf{T}}_j = \mathbf{Y}_j \bar{\mathbf{P}} \quad (4)$$

where $\mathbf{Y}_j$ is the j -th observation of $\mathbf{Y}$, $\bar{\mathbf{T}}_j$ is the corresponding row of $\bar{\mathbf{T}}$, and $\bar{\mathbf{P}}$ is the loading matrix that contains the coefficients for the transformation.

4.2 NCA for FS.

NCA has been applied to reduce dimensionality in AR [52,53]. The objective is to find a vector of weighting, $\mathbf{w}$, to select the subset of characteristics optimizing the classification of the nearest neighbours, being $\mathbf{Y} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n\}$ the labelled features set, $\mathbf{y}_i \in \mathcal{R}^d$, and the corresponding class label are $\{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\}$. Weighting vector $\mathbf{w}$ is the weighted distance between two samples $\mathbf{y}_i$ and $\mathbf{y}_j$, (equation (5)):

$$D_{\mathbf{w}}(\mathbf{y}_i, \mathbf{y}_j) = \sum_{l=1}^d \mathbf{w}_l^2 |\mathbf{y}_{il} - \mathbf{y}_{jl}| \quad (5)$$

where $\mathbf{w}_l$ is a weight associated with l -th feature.

The reference point is chosen randomly in NCA, and all points in $\mathbf{X}$ are likely to be selected as the reference point. The point $\mathbf{x}_i$ selects the point $\mathbf{x}_j$ as its neighbour with the probability p_{ij} ,

receiving the class label from the neighbour selected. The probability p_{ij} is defined by equation (6):

$$\begin{cases} p_{ij} = \frac{\mathcal{K}(D_{\mathbf{w}}(\mathbf{y}_i, \mathbf{y}_j))}{\sum_{k \neq i} \mathcal{K}(D_{\mathbf{w}}(\mathbf{y}_i, \mathbf{y}_k))}; \text{ if } i \neq j \\ p_{ij} = 0; \text{ if } i = j \end{cases} \quad (6)$$

where $\mathcal{K}(z) = \exp(-\frac{z}{\sigma})$ is a kernel function, and σ is an input parameter that set the probability of each point selected as the reference point. The probability of the query point $\mathbf{x}_i$ correctly classified is given by equation (7):

$$p_i = \sum_j C_{ij} p_{ij} \quad (7)$$

where $C_{ij} = 1$ if the class label is $C_i = C_j$, and $C_{ij} = 0$ if the class label is $C_i \neq C_j$.

The true leave-one-out classification accuracy $\xi(w)$ is calculated by equation (8):

$$\xi(w) = \sum_i \sum_j C_{ij} p_{ij} - \lambda \sum_{l=1}^d w_l^2 \quad (8)$$

where λ is a regularization parameter set by cross validation. The regularization term $\lambda \sum_{l=1}^d w_l^2$ is introduced in the equation (8) to improve the feature selection and avoid overfitting.

A gradient rule is applied by differentiating $\xi(w)$ regarding w_l in equation (9):

$$\frac{\partial \xi(w)}{\partial w_l} = 2 \left(\frac{1}{\sigma} \sum_i \left(p_i \sum_{j \neq i} p_{ij} |y_{il} - y_{jl}| - \sum_j C_{ij} p_{ij} |y_{il} - y_{jl}| - \lambda \right) w_l \right) \quad (9)$$

4.3 Classifiers

Two scenarios have been considered to test the accuracy of the classifiers:

- **E-1.** Classification levels with short ultrasonic pulses at the same frequency (25 kHz).
- **E-2.** Classification levels with short ultrasonic pulses at different frequencies [18, 20, 25, 30, 40, 50] kHz.

4.3.1 Machine learning

Decision Trees (DT).

DT is a classifier used to determine if the dataset contains different classes of objects that can be interpreted significantly in the context of a substantive theory [54]. DT predicts responses to the dataset following the decisions in the tree from the root node (beginning) down to a leaf node. The leaf node contains the binary response. Binary term, except for the terminal node, will be

divided into two secondary nodes, based on Splitting Rules (SR). The purpose of the SR is to minimize the impurity of the node. In this case, SR is based on the Gini's Diversity Index (GDI) [55], given by equation (10).

$$GDI = 1 - \sum_i p^2(i) \quad (10)$$

where the sum is over the classes i at the node, and $p(i)$ is the observed fraction of classes with class i that has the node. A node with only one class, called a pure node, $GDI = 0$, being $GDI > 0$ in other cases. GDI is therefore a measure of node impurity.

SR can be modified by cost function to adjust the classification error, or to compensate for the unequal sampling of classes [56]. DT grows until there is no possibility of creating additional splits, or when the process stops because the predefined objective has been achieved. The values of the terminal nodes are assigned to each class of mud level.

The fine, medium and coarse tree models of DT have been tested, where the main difference is the maximum number of splits, being 100 for the fine, 20 for the medium and 4 for the coarse tree model. The index split in all cases has been set by the GDI criterion for setting. The cross-validation technique [57] has been used against overfitting and stimulating accuracy in each fold, where 5 folds have been applied according to the tests.

Linear Discriminant Analysis (LDA).

The ultrasonic signals are homogeneous for the same frequency. This is the main reason that the LDA classifier provides better results [58]. This classifier is also valid for ultrasonic signals obtained at different frequencies. It is based on a geometrical approach, where a linear function, representing a decision limit, divides a feature space into regions that have common properties. The aim of this method is to find the expression of linear discriminant functions that enable the classification of objects in the considered classes. The linear discriminant function $g(x)$ is given by equation (11)

$$g(x) = w_0 + \sum_{i=1}^d w_i x_i \quad (11)$$

where the coefficients w_i are the components of the weight vector $\mathbf{w}$.

The model has the same covariance matrix for each class, and only the means vary. The classifier calculates the sample mean of each class, then it computes the sample covariance. The hyperparameters are set by linear coefficient threshold (Delta) and coefficient of regularization (Gamma) to minimize the loss of five-fold cross validation.

Linear Support Vector Machine (LSVM)

LSVM is a supervised multivariate classification method, where “supervised” refers to a training step in which the algorithm learns the differences between pre-specified groups to be classified [59]. LSVM treats each feature as a point in a high-dimensional space, being the number of dimensions equal to the number of rating levels. Each feature is assigned to a group, and the linear classification function learns the characteristics that must be discriminated into the five groups. A limit of decision, or hyperplane (a generalization of a plane of $n-1$ dimensions which divides an n -dimensional space), is defined to cluster the dataset, based on class membership to classify the dataset linearly. However, for a linearly separable problem there are infinite hyperplanes correctly classified for the dataset. LSVM algorithm finds the optimal solution characterized by the largest margin between classes. The margin is defined as the distance of the closest training data points of the hyperplane. These points are the most difficult to classify and are called support vectors. The hyperplane is defined by a weight vector, which is a linear combination of the support vectors, and specifies a direction and a displacement that define the maximum margin classifier.

The decision function is given by equation (12),

$$D(x) = w\phi(x) + b \quad (12)$$

where w and b are the SVM parameters, and $\phi(x)$ is a kernel function.

The hyperplane is defined by equation (12), and the distance between the hyperplane and pattern x is written as $\frac{D(x)}{\|w\|}$.

A training classifier is designed to find the value w that maximizes the difference between the class boundary and the training patterns. The objective function of the training algorithm is given by equation (13).

$$J = \frac{\|w\|^2}{2} \quad (13)$$

The Linear, Polynomial and Gaussian Radial Basis Kernel functions [54] have been used to modify the different Kernel parameters in each case. The Kernel linear function provided the best results. The Sequential Minimal Optimization (SMO) method has been applied to set the training algorithm to find the separating hyperplane.

The tolerance and the fraction of variables allowed to violate Karush-Kuhn-Tucker [55] (KKT) conditions have also been adjusted for the SMO training method. The algorithm automatically centres the data points at their mean and scales them to a unit standard deviation before training.

The box constraint level parameter controls the maximum penalty imposed on observations that violate the soft margin and helps to prevent excessive adjustment (regulation). The best box constraint level parameter found in this paper is 1. The Kernel scale parameter is selected based on a heuristic procedure.

One against one method has been implemented for multiclass classification [60]. All possible two-class classifiers are evaluated from the training set of n classes, each classifier being trained in only two out of n classes, giving a total of $n(n-1)/2$ classifiers. The label of the class with most votes is assigned to the dataset, where a vote is given to the winning class, applying each classifier to the test dataset.

k-Nearest Neighbours (k-NN)

The k-NN is a high-performance classifier used in ML. k-NN rule classification is an extension of the Nearest-Neighbour (NN) rule. It assigns to each sample the label of the most frequently represented over the k nearest samples [61]. The relative simplicity of the k-NN search technique means that the results from other classification techniques can be compared.

The accuracy of k-NN classification depends on the metric used to compute distances between different samples. In this case, the best performing classifier is Fine k-NN (FKNN) [62], using metric Euclidean distance given by equation (14).

$$d_{st}^2 = (x_s - y_t) \cdot (x_s - y_t)' \quad (14)$$

The number of neighbours, k , affects the classification performance of the Fine k-NN algorithm [63]. A larger k could be used when noise is presented in the locality of the query instance. A smaller k is employed when the region defining the class is small. In this approach, different numbers of k (0-10) have been tested. The best result is obtained with a value of $k=5$.

Ensemble Subspace Discriminant (ESD)

ESD is a predictive model composed of a weighted combination of multiple classification models. Bagging, boosting and the random subspace methods are employed to improve weak classifiers. In general, combining multiple classification models increases predictive performance. The Random Subspace Method (RSM) is a parallel learning algorithm proposed by Ho [64], where the generation of a discriminant classifier is independent in each subspace from the original dimensional feature space. The training data is then modified and performed in the feature space [65]. The same hyperparameters have been used for LDA.

4.3.2 Evaluation of classifiers

The confusion matrix (Table 2) is employed to evaluate the classifiers by the Receiver Operating Characteristic (ROC) analysis. The confusion matrix determines the accuracy of a classifier and measures its performance. Table 2 presents the confusion matrix compounds by the true positive, TP , the real successes of the classifier, and the error, E .

TABLE 2
CONFUSION MATRIX

ACTUAL	PREDICTED					
	Level	1	2	3	4	5
	1	TP	E ₁₂	E ₁₃	E ₁₄	E ₁₅
	2	E ₂₁	TP	E ₂₃	E ₂₄	E ₂₅
	3	E ₃₁	E ₃₂	TP	E ₃₄	E ₃₅
	4	E ₄₁	E ₄₂	E ₄₃	TP	E ₄₅
	5	E ₅₁	E ₅₂	E ₅₃	E ₅₄	TP

The confusion matrix has the following characteristics:

- The total number of test examples of any class is the sum of the corresponding row.
- The total number of false negatives (*FN*) for a class is the sum of values in the corresponding row, excluding the *TP*.
- The total number of false positives (*FP*) for a class is the sum of values in the corresponding column, excluding the *TP*.
- The total number of true negatives (*TN*) for a certain class is the sum of all columns and rows, excluding the column and row of the class.

The main parameters obtained from the confusion matrix are:

- Recall is the probability of being correctly classified, known as the true positive rate (equation 15):

$$Recall = \frac{TP}{TP + FN} \quad (15)$$

- The true positives rate is given by equation (16):

$$TP\ rate \approx \frac{Positives\ correctly\ classified}{Total\ Positives} \quad (16)$$

- The false positive rate, called the false alarm rate, of the classifier (equation (17)):

$$FP\ rate \approx \frac{Negatives\ incorrectly\ classified}{Total\ negatives} \quad (17)$$

- The true negative rate, that measures the proportion of negatives that are correctly identified (equation (18)):

$$Specificity\ (TN\ rate) = \frac{TN}{FP + TN} \quad (18)$$

Additional terms associated with ROC curves and confusion matrix are:

- Precision (equation (19)):

$$Precision = \frac{TP}{FP + TP} \quad (19)$$

- Accuracy (equation (20)):

$$Accuracy = \frac{TP + TN}{P + N} \quad (20)$$

- F-score (equation (21)):

$$F - score = \frac{2 \times precision \times recall}{precision + recall} \quad (21)$$

There are two conventional methods to establish the average performance in all categories: macro-averaging and micro-averaging [66].

- *Macro-average* (P_i^M, R_i^M, F_i^M): P^M, R^M, F^M are obtained by the averaging over all P_i^M, R_i^M, F_i^M , where M denotes Macro-average and i is the scenario. They are calculated for each scenario, i.e. the precision P , recall R and F-score F values are evaluated locally P_i^M, R_i^M, F_i^M , and then globally P^M, R^M, F^M .
- *Micro-average* ($P_i^\mu, R_i^\mu, F_i^\mu$): P, R and F are obtained to micro-average, μ , as: TP_i, FP_i and FN_i . The value of TP, FP and FN are calculated as the sum of TP_i, FP_i and FN_i .

Area Under Curve (AUC) is the area between the ROC and the negative diagonal [67]. The value is between 0.5 and 1, where the value 0.5 indicates that the classifier is not valid because the TN rate is the same as the FP . The AUC of a classifier is equivalent to the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance, i.e. it is equivalent to the Wilcoxon test of ranks [67]. The AUC is also related to the Gini coefficient [68], being twice the area between the diagonal and the ROC curve. AUC is given by equation (22):

$$AUC = \frac{TP\ rate - FP\ rate + 1}{2} \quad (22)$$

The recommendations by Demšar [69], together with the extensions by Garcia and Herrera [70], have been employed for a comparative analysis of the classifiers. Friedman test [71] has been used to study the null hypothesis so that all classifiers obtain the same average. Bonferroni-Dunn Test [72] has been applied to determine significant differences between the top-ranked classifier and the next in the rank. Finally, Holm Test [73] has been done to corroborate the results.

The t-Test [74] has also been applied to confirm the null hypothesis and validate the results.

5 Results

5.1 FE

The NCA has been applied for FS. NCA selects 5 features for the E1 scenario from the 15 features extracted by the AR method, (see Figure 8). It has been applied to avoid the so called “course of dimensionality”.

Figure 8. NCA for E-1.

5.2 Measure precision

Table 3 shows the experimental Precision results provided by the 5 classifiers for E-1 (signal at 25 kHz). The first column indicates the level of mud and the precision parameters. The last two rows show the ranking by the classifiers and their position given by Friedman tests. The table is divided into five main columns corresponding to the classifiers, which are subdivided into two sub-columns. The first sub-column shows the precision results by the AR FE/NCA FS methods, and the second corresponds to the PCA method. The values of micro and macro precisions are included in the sixth and seventh rows.

TABLE 3
RESULTS PRECISION P_i (E-1) (25 kHz)

LEVEL OF MUD	DT		LDA		LSVM		FKNN		ESD	
	AR	PCA	AR	PCA	AR	PCA	AR	PCA	AR	PCA
1	0.50	1.00	0.50	1.00	0.90	1.00	0.60	1.00	0.90	1.00
2	0.80	0.90	0.80	1.00	0.90	1.00	0.90	1.00	0.90	1.00
3	0.80	1.00	0.90	1.00	0.80	1.00	0.80	1.00	0.90	1.00
4	0.70	1.00	0.60	1.00	0.90	1.00	0.60	1.00	0.90	1.00
5	0.80	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
P^{μ}	0.72	0.98	0.76	1.00	0.90	1.00	0.78	1.00	0.92	1.00
P^M	0.717	0.982	0.768	1.00	0.916	1.00	0.84	1.00	0.93	1.00
RANKING	1.80	2.60	2.50	3.10	3.70	3.10	2.80	3.10	4.20	3.10
CLASSIFIC.	5	1	4	1	2	1	3	1	1	1

Table 3 shows the precision results for E-1. Level 1, for AR, provides lower precision, with 50% in the worst case for DT, and 90% in the best case, where level 5 shows the best results. The results of micro-average are from 72% to 92%, and the macro-average from 71.7% to 93 %. The PCA method shows better precision values in most classifiers.

Table 4 contains the precision results for E-2. The table shows that level 1 provides the lower precision for the AR FE method, 38% in the worst case, and 65% in the best, where level 5 gives the best results. The results of micro-average are from 55% to 65%, and micro-average from 54.9% to 65.9%. PCA shows excellent and poor results.

The significant difference between the classifiers studied has been analysed according to Demšar's recommendations. The test indices indicate the comparisons of the different classifiers shown in Tables 5-7 (first column) (TEST), where 1 is DT, 2 - LDA, 3 - LSVM, 4 - FKNN and 5 - ESD. According to the Friedman test for E-1, AR rejects the null hypothesis because no significant differences were found ($p\text{-value} = 0.0471 \leq 0.05$). Holm test (Table 5) rejects the null hypothesis for DT and ESD classifiers.

PCA did not reject the null hypothesis for E-1 ($p\text{-value} = 0.4060 \geq 0.05$). The Bonferroni-Dunn test did not reject the null hypothesis for $p \leq 0.05$, with $\alpha = 0.05$ (Figure 9).

TABLE 4
RESULTS PRECISION P_i (E-2)

LEVEL OF MUD	DT		LDA		LSVM		FKNN		ESD	
	AR	PCA	AR	PCA	AR	PCA	AR	PCA	AR	PCA
1	0.55	0.92	0.38	0.33	0.65	0.38	0.65	1.00	0.67	0.33
2	0.72	0.97	0.48	0.28	0.68	0.55	0.77	1.00	0.63	0.30
3	0.30	0.93	0.70	0.22	0.62	0.54	0.68	1.00	0.55	0.18
4	0.37	0.90	0.33	0.10	0.45	0.63	0.68	1.00	0.48	0.10
5	0.83	1.00	0.63	0.43	0.83	0.50	0.90	1.00	0.88	0.50
P^H	0.55	0.94	0.51	0.27	0.65	0.52	0.74	1.00	0.64	0.28
P^M	0.549	0.94	0.527	0.27	0.659	0.52	0.739	1.00	0.651	0.28
RANKING	2,30	4,10	1,80	1,40	3,00	2,90	4,50	4,90	3,40	1,70
CLASSIFIC.	4	2	5	5	3	3	1	1	2	4

AR did not reject for E-2 ($p\text{-value} = 0.4060 \geq 0.05$), while PCA rejects the null hypothesis for E-2 ($p\text{-value} = 0.000816 \leq 0.05$). Holm test rejects the null hypothesis for DT and ESD classifiers.

TABLE 5.
RESULTS TEST HOLM.

Test	p-value	alpha	Comment
2-4	0.0000	0.0051	Reject Ho
4-5	0.0000	0.0057	Reject Ho
1-2	0.0000	0.0064	Reject Ho
1-5	0.0000	0.0073	Reject Ho
3-4	0.0000	0.0085	Reject Ho
1-3	0.0000	0.0102	Reject Ho
2-3	0.0007	0.0127	Reject Ho
3-5	0.0011	0.0170	Reject Ho
1-4	0.3797	0.0253	Fail to reject Ho
3-5	0.8742	No comparison	Ho is accepted

Figure 9. Measure precision Bonferroni-Dunn test for PCA in E-2. The mean column rank of LDA, ESD and FKNN groups are significantly different.

The results for E-1 indicate that ESD is the most accurate classifier (Table 3), and FKNN is the best for E-2 (Table 4). The test indicates that LDA can be discarded for all cases. ESD cannot be considered for E-2 either.

When the t-Test is applied for AR and E-1 (Table 6), the test rejects the null hypothesis in the cases of the 1- 3 and 1-5 classifiers. Consequently, classifier 1, DT can be discarded. The T-test does not reject the null hypothesis for PCA and E-1.

**TABLE 6.
RESULTS T-TEST: AR AND E-1.**

Test	p-value	Comment
1-3	0.0265	Reject Ho
1-5	0.0118	Reject Ho

Table 7 shows t-Test results for AR and E-2. LDA classifier is discarded.

**TABLE 7.
RESULTS T-TEST: AR AND E-2.**

Test	p-value	Comment
2-4	0.0250	Reject Ho

5.3 Measure recall

Tables 8 and 9 show the Recall results obtained from the E-1 and E-2 experiments. Bonferroni-Dunn test rejects the null hypothesis for DT and ESD. Friedman test does not reject the null hypothesis for PCA and E-1 ($p\text{-value} = 0.3504 \geq 0.05$). Friedman test rejects the null hypothesis

for E-2 ($p\text{-value} = 0.0061 \leq 0.05$; $p\text{-value} = 0.0014 \leq 0.05$), according to the null hypothesis tested with post-hoc test. Bonferroni-Dunn test rejects the null hypothesis for DT, LDA and FKNN. The t-Test confirms these results.

TABLE 8
RESULTS RECALL R_i (E-1)

LEVEL OF MUD	DT		LDA		LSVM		FKNN		ESD	
	AR	PCA	AR	PCA	AR	PCA	AR	PCA	AR	PCA
1	0.63	1.00	0.50	1.00	0.90	0.91	0.86	1.00	1.00	1.00
2	0.89	1.00	0.89	1.00	1.00	1.00	0.90	1.00	1.00	1.00
3	0.80	1.00	0.64	1.00	0.89	1.00	0.62	0.90	0.90	1.00
4	0.54	0.91	0.71	1.00	0.75	1.00	0.60	1.00	0.75	1.00
5	0.80	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
R^μ	0.72	0.98	0.74	1.00	0.90	0.98	0.78	0.98	0.92	1.00
R^M	0.717	0.98	0.74	1.00	0.90	0.98	0.78	0.98	0.92	1.00
RANKING	1.70	2.40	2.20	3.50	4.10	3.00	2.50	2.60	4.50	3.50
CLASSIFIC.	5	5	4	2	2	3	3	4	1	1

TABLE 9
RESULTS RECALL R_i (E-2)

LEVEL OF MUD	DT		LDA		LSVM		FKNN		ESD	
	AR	PCA	AR	PCA	AR	PCA	AR	PCA	AR	PCA
1	0.52	0.95	0.46	0.45	0.61	0.44	0.67	1.00	0.60	0.45
2	0.62	0.85	0.63	0.15	0.77	0.39	0.74	1.00	0.81	0.15
3	0.56	1.00	0.43	0.31	0.53	0.77	0.71	1.00	0.57	0.31
4	0.35	0.95	0.36	0.15	0.46	0.45	0.63	1.00	0.45	0.15
5	0.69	0.98	0.76	0.45	0.93	0.75	0.95	1.00	0.83	0.48
R^μ	0.55	0.94	0.51	0.27	0.65	0.52	0.74	1.00	0.64	0.28
R^M	0.55	0.94	0.51	0.27	0.65	0.52	0.74	1.00	0.64	0.28
RANKING	0.60	4.10	1.60	1.60	3.60	2.60	4.60	4.90	3.60	1.80
CLASSIFIC.	4	2	5	5	2	3	1	1	3	4

Friedman test rejects the null hypothesis for AR and E-1 ($p\text{-value} = 0.0077 \leq 0.05$), according to the null hypothesis tested with post-hoc test. Figure 10 shows the results of the Bonferroni test.

Figure 10. Measure Recall Bonferroni-Dunn test for E-1

5.4 Measure F-Score.

Tables 10 and 11 provide the F-score results for E-1 and E-2.

Friedman test rejects the null hypothesis for AR and E-1 ($p\text{-value} = 0.0081 \leq 0.05$). The post-hoc test indicates that there is a significant difference between LDA and 5 ESD for AR and E-1. Friedman test ($p\text{-value} = 0.3084 \geq 0.05$) does not reject the null hypothesis for PCA and E-1.

The classifier with better accuracy for AR and E-2 is KFNN, and the Bonferroni test discard classifiers DT for the AR case, and LDA and ESD for PCA. The Holm test and T-test confirm these results.

TABLE 10
RESULTS F-SCORE F_i (E-1)

LEVEL OF MUD	DT		LDA		LSVM		FKNN		ESD	
	AR	PCA	AR	PCA	AR	PCA	AR	PCA	AR	PCA
1	0.56	1.00	0.50	1.00	0.9	0.95	0.71	1.00	0.95	1.00
2	0.84	0.95	0.84	1.00	0.95	1.00	0.90	0.95	0.95	1.00
3	0.80	1.00	0.75	1.00	0.84	0.95	0.70	0.95	0.90	1.00
4	0.61	0.95	0.59	1.00	0.75	1.00	0.60	1.00	0.82	1.00
5	0.80	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
F^u	0.72	0.98	0.74	1.00	0.90	0.98	0.78	0.98	0.92	1.00
F^M	0.72	0.98	0.74	1.00	0.90	0.98	0.78	0.98	0.92	1.00
RANKING	2,10	2,60	1,80	3,60	4	2,60	2,50	2,60	4,60	3,60
CLASSIFIC.	4	5	5	2	2	3	3	4	1	1

TABLE 11
RESULTS F-SCORE F_i (E-2)

LEVEL OF MUD	DT		LDA		LSVM		FKNN		ESD	
	AR	PCA	AR	PCA	AR	PCA	AR	PCA	AR	PCA
1	0.53	0.99	0.42	0.38	0.63	0.41	0.66	1.00	0.63	0.38
2	0.67	0.96	0.55	0.19	0.73	0.46	0.75	1.00	0.71	0.20
3	0.39	1.00	0.53	0.25	0.57	0.64	0.69	1.00	0.56	0.23
4	0.36	0.99	0.34	0.12	0.45	0.53	0.66	1.00	0.47	0.12
5	0.76	1.00	0.69	0.44	0.88	0.60	0.92	1.00	0.85	0.49
F^H	0.54	0.94	0.51	0.28	0.65	0.53	0.74	1.00	0.64	0.29
F^M	0,54	0,94	0,51	0,28	0,65	0,53	0,74	1,00	0,64	0,29
RANKING	1,80	4,20	1,20	1,40	3,70	3,00	5	4,80	3,30	1,60
CLASSIFIC.	4	2	5	5	2	3	1	1	3	5

5.5 AUC results.

Tables 12 and 13 present AUC results for E-1 and E-2. Friedman and Post-Hoc tests exclude the DT classifier for AR and E-1. DT and LDA are excluded for E-2.

TABLE 12
RESULTS AUC (E-1)

LEVEL OF MUD	DT		LDA		LSVM		FKNN		ESD	
	AR	PCA	AR	PCA	AR	PCA	AR	PCA	AR	PCA
1	0.80	1.00	0.75	1.00	0.94	1.00	0.90	1.00	0.98	1.00
2	0.89	0.99	0.93	1.00	0.99	1.00	0.94	1.00	0.99	1.00
3	0.85	1.00	0.82	1.00	0.93	0.99	0.89	1.00	0.94	1.00
4	0.81	0.95	0.83	1.00	0.87	1.00	0.75	1.00	0.86	1.00
5	0.89	1.00	1.00	1.00	1.00	1.00	0.95	1.00	1.00	1.00
RANKING	1.60	2.30	2.20	3.30	4.30	2.80	2.40	3.30	4.50	3.30
CLASSIFIC.	5	3	4	1	2	2	3	1	1	1

The classifier that detects and identifies mud in all cases is ESD for E-1, with an accuracy of 93% (AR) and 100% (PCA). FKNN is the best classifier for E-2, with a precision of 74% (AR) and 100% (PCA). The most effective classifier is FKNN for PCA in all cases. Level 5 (no mud) is usually detected and identified with high accuracy, considering the AUC parameters (Tables 12 and 13), with a minimal success rate of 89% (E-1) and 67% (E-2). The remaining mud levels

demonstrate a success rate of around 60%. This indicates the robustness of the classifier for any frequency, especially for the FKNN classifier.

TABLE 13
RESULTS AUC (E-2)

LEVEL OF MUD	DT		LDA		LSVM		FKNN		ESD	
	AR	PCA	AR	PCA	AR	PCA	AR	PCA	AR	PCA
1	0.76	0.96	0.74	0.66	0.79	0.69	0.82	1.00	0.79	0.63
2	0.80	0.97	0.80	0.62	0.87	0.73	0.86	1.00	0.88	0.62
3	0.77	0.99	0.72	0.59	0.76	0.76	0.84	1.00	0.78	0.60
4	0.70	0.95	0.70	0.64	0.73	0.81	0.80	1.00	0.73	0.64
5	0.84	1.00	0.86	0.71	0.95	0.75	0.96	1.00	0.90	0.67
RANKING	1,80	4,10	1,40	1,60	3,40	3,00	4,60	4,90	3,80	1,40
CLASSIFIC.	4	2	5	4	3	3	1	1	2	5

6 Conclusions

This work discusses an approach that detects and diagnoses the levels of mud on wind turbine blades by guided waves. The approach filters the dataset by Wavelet transform, and the dimension of the signal is reduced by Feature Extraction and Feature Selection. Neighbourhood Component Analysis has been employed in Feature Selection. Pattern recognition is carried out by supervised learning classifiers (Ensemble Subspace Discriminant (ESD), k-Nearest Neighbours (kNN, Linear Support Vector Machine (LSVM), Linear Discriminant Analysis (LDA) and Decision Trees (DT)) based on statistical models to calculate and classify the faults. Autoregressive (AR) model and Principal Component Analysis (PCA) have been employed to Feature Extraction. Receiver operating characteristic analysis is used to evaluate the classifiers.

Two scenarios have been studied: single ultrasonic frequency of the guided waves at 25 kHz (E-1), and multi-frequency analysis at 18, 20, 25, 30 40 and 50 kHz (E-2). Different amounts of mud have been placed on an actual blade section to test the effectiveness of the presented techniques.

The results show that the combination of the KNN with the PCA provide the best results for the detection and diagnosis of mud in the experiments considered. The classifier that detects and identifies mud in all cases is ESD for E-1. FKNN is the best classifier for E-2.

The methodology detects and classifies the levels of mud considered in the case study based on a real wind turbine blade section.

Acknowledgements

The work reported herewith has been financially supported by the Spanish Ministerio de Economía y Competitividad, under Research Grants DPI2015-67264-P and RTC-2016-5694-3.

7 References

1. Gómez Muñoz, C.Q.; Villegas, M.A.; García, F.P.; Pedregal, D.J. Big data and web intelligence for condition monitoring: A case study on wind turbines. *Handbook of Research on Trends and Future Directions in Big Data and Web Intelligence; Information Science Reference, IGI Global: Hershey, PA, USA* **2015**.
2. Gómez Muñoz, C.Q.; Ruiz de la Hermosa Gonzalez-Carrato, R.; Trapero Arenas, J.R.; García Marquez, F.P. A novel approach to fault detection and diagnosis on wind turbines. *GLOBAL NEST JOURNAL* **2014**, *16*, 1029-1037.
3. Simola, K.; Pulkkinen, U. Models for non-destructive inspection data. *Reliability Engineering & System Safety* **1998**, *60*, 1-12.
4. Dong, W.; Moan, T.; Gao, Z. Fatigue reliability analysis of the jacket support structure for offshore wind turbine considering the effect of corrosion and inspection. *Reliability Engineering & System Safety* **2012**, *106*, 11-27.
5. Papaelias, M.; Cheng, L.; Kogia, M.; Mohimi, A.; Kappatos, V.; Selcuk, C.; Constantinou, L.; Gómez Muñoz, C.Q.; Marquez, F.P.G.; Gan, T.-H. Inspection and structural health monitoring techniques for concentrated solar power plants. *Renewable Energy* **2016**, *85*, 1178-1191.
6. Pliego Marugán, A.; García Márquez, F.; Pinar Pérez, J. Optimal maintenance management of offshore wind farms. *Energies* **2016**, *9*, 46.
7. García Márquez, F.P.; Pliego Marugán, A.; Pinar Pérez, J.M.; Hillmansen, S.; Papaelias, M. Optimal dynamic analysis of electrical/electronic components in wind turbines. *Energies* **2017**, *10*, 1111.
8. Pérez, J.M.P.; Márquez, F.P.G.; Hernández, D.R. Economic viability analysis for icing blades detection in wind turbines. *Journal of Cleaner Production* **2016**, *135*, 1150-1160.
9. Rodriguez, Z.; Alonso, O. Soiling shield for wind turbine blade. US Patent 20,150,322,916: 2015.
10. BladeCleaning. [Http://www.BladeCleaning.Com/problematica_en.Htm](http://www.BladeCleaning.Com/problematica_en.Htm). **2010**.
11. Thresher, R.W.; Robinson, M.C.; Veers, P.S. *Wind energy technology: Current status and r & d future*. National Renewable Energy Laboratory: 2008.
12. Tangler, J.L.; Somers, D.M. Airfoils for wind turbine. Google Patents: 2000.
13. Van Dam, C.; Mayda, E.; Chao, D.; Jackson, K.; Zuteck, M.; Berry, D. In *Innovative structural and aerodynamic design approaches for large wind turbine blades*, 2005 ASME Wind Energy Symposium/The 43th AIAA Aerospace Sciences Meeting & Exhibit. Reno, NV, 2005; pp 68-98.
14. Gómez Muñoz, C.Q.; García Marquez, F.P.; Liang, C.; Maria, K.; Abbas, M.; Mayorkinos, P. A new condition monitoring approach for maintenance management in concentrate solar plants. In *Proceedings of the ninth international conference on management science and engineering management*, Springer Berlin Heidelberg: 2015; pp 999-1008.
15. Márquez, F.P.G.; Pedregal, D.J.; Roberts, C. New methods for the condition monitoring of level crossings. *International Journal of Systems Science* **2015**, *46*, 878-884.
16. Márquez, F.P.G.; Pardo, I.P.G.; Nieto, M.R.M. Competitiveness based on logistic management: A real case study. *Annals of Operations Research* **2013**, 1-13.
17. Jiménez, A.A.; Muñoz, C.Q.G.; Marquez, F.P.G.; Zhang, L. In *Artificial intelligence for concentrated solar plant maintenance management*, Proceedings of the Tenth International Conference on Management Science and Engineering Management, 2017; Springer: pp 125-134.
18. Muñoz, C.Q.G.; Marquez, F.P.G.; Liang, A.A.J.; Cheng, M.K.; Mohimi, A.; Papaelias, M.; Mohimi, A.; Cheng, L. Fault detection and diagnosis employing the electromagnetic sensors emat.

19. Niu, G.; Yang, B.-S.; Pecht, M. Development of an optimized condition-based maintenance system by data fusion and reliability-centered maintenance. *Reliability Engineering & System Safety* **2010**, *95*, 786-796.
20. Gómez Muñoz, C.Q.; García Márquez, F.P. A new fault location approach for acoustic emission techniques in wind turbines. *Energies* **2016**, *9*, 40.
21. Márquez, F.P.G.; Pérez, J.M.P.; Marugán, A.P.; Papaelias, M. Identification of critical components of wind turbines using fta over the time. *Renewable Energy* **2015**.
22. Muñoz, C.Q.G.; Márquez, F.P.G.; Tomás, J.M.S. Ice detection using thermal infrared radiometry on wind turbine blades. *Measurement* **2016**, *93*, 157-163.
23. Liu, J.; Zio, E. System dynamic reliability assessment and failure prognostics. *Reliability Engineering & System Safety* **2017**, *160*, 21-36.
24. de la Hermosa González, R.R.; Márquez, F.P.G.; Dimlaye, V.; Ruiz-Hernández, D. Pattern recognition by wavelet transforms using macro fibre composites transducers. *Mechanical Systems and Signal Processing* **2014**, *48*, 339-350.
25. de la Hermosa González, R.R.; Márquez, F.P.G.; Dimlaye, V. Maintenance management of wind turbines structures via mfcs and wavelet transforms. *Renewable and Sustainable Energy Reviews* **2015**, *48*, 472-482.
26. Marks, R.; Clarke, A.; Featherston, C.; Paget, C.; Pullin, R. Lamb wave interaction with adhesively bonded stiffeners and disbonds using 3d vibrometry. *Applied Sciences* **2016**, *6*, 12.
27. Muñoz, C.Q.G.; Marquez, F.P.G.; Lev, B.; Arcos, A. New pipe notch detection and location method for short distances employing ultrasonic guided waves. *Acta Acustica united with Acustica* **2017**, *103*, 772-781.
28. Farrar, C.R.; Doebling, S.W.; Nix, D.A. Vibration-based structural damage identification. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **2001**, *359*, 131-149.
29. Muñoz, C.Q.G.; Jiménez, A.A.; Márquez, F.P.G. Wavelet transforms and pattern recognition on ultrasonic guides waves for frozen surface state diagnosis. *Renewable Energy* **2017**.
30. Staszewski, W.J.; Robertson, A.N. Time-frequency and time-scale analyses for structural health monitoring. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **2007**, *365*, 449-477.
31. Martinez-Luengo, M.; Kolios, A.; Wang, L. Structural health monitoring of offshore wind turbines: A review through the statistical pattern recognition paradigm. *Renewable and Sustainable Energy Reviews* **2016**, *64*, 91-105.
32. Cherkassky, V.; Mulier, F.M. *Learning from data: Concepts, theory, and methods*. John Wiley & Sons: 2007.
33. Pliego Marugán, A.; García Márquez, F.P.; Lev, B. Optimal decision-making via binary decision diagrams for investments under a risky environment. *International Journal of Production Research* **2017**, *1-16*.
34. Hameed, Z.; Hong, Y.; Cho, Y.; Ahn, S.; Song, C. Condition monitoring and fault detection of wind turbines and related algorithms: A review. *Renewable and Sustainable energy reviews* **2009**, *13*, 1-39.
35. Rankin, A.L.; Matthies, L.H. Passive sensor evaluation for unmanned ground vehicle mud detection. *Journal of Field Robotics* **2010**, *27*, 473-490.
36. Milinda, H.; Madhusanka, B. In *Mud and dirt separation method for floor cleaning robot*, Engineering Research Conference (MERCon), 2017 Moratuwa, 2017; IEEE: pp 316-320.
37. Rajpal, P.; Shishodia, K.; Sekhon, G. An artificial neural network for modeling reliability, availability and maintainability of a repairable system. *Reliability Engineering & System Safety* **2006**, *91*, 809-819.
38. Pérez-Ortiz, M.; Jiménez-Fernández, S.; Gutiérrez, P.A.; Alexandre, E.; Hervás-Martínez, C.; Salcedo-Sanz, S. A review of classification problems and algorithms in renewable energy applications. *Energies* **2016**, *9*, 607.
39. Dervilis, N.; Choi, M.; Taylor, S.; Barthorpe, R.; Park, G.; Farrar, C.; Worden, K. On damage diagnosis for a wind turbine blade using pattern recognition. *Journal of sound and vibration* **2014**, *333*, 1833-1850.
40. Pavlopoulou, S.; Worden, K.; Soutis, C. Novelty detection and dimension reduction via guided ultrasonic waves: Damage monitoring of scarf repairs in composite laminates. *Journal of Intelligent Material Systems and Structures* **2016**, *27*, 549-566.

41. Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. *Chemometrics and intelligent laboratory systems* **1987**, 2, 37-52.
42. de Lautour, O.R.; Omenzetter, P. Damage classification and estimation in experimental structures using time series analysis and pattern recognition. *Mechanical Systems and Signal Processing* **2010**, 24, 1556-1569.
43. Jolliffe, I. *Principal component analysis*. Wiley Online Library: 2002.
44. Hunter, N.F.; Worden, K. Using statistical pattern recognition techniques. *Engineering Analysis* **2001**, 1000, C926.
45. Zhang, T.; Chen, W.; Li, M. Ar based quadratic feature extraction in the vmd domain for the automated seizure detection of eeg using random forest classifier. *Biomedical Signal Processing and Control* **2017**, 31, 550-559.
46. Gómez Muñoz, C.Q.; Arcos Jimenez, A.; García Marquez, F.P.; Kogia, M.; Cheng, L.; Mohimi, A.; Papaelias, M. Cracks and welds detection approach in solar receiver tubes employing electromagnetic acoustic transducers. *Structural Health Monitoring* **2017**, 1475921717734501.
47. Eristi, H. Fault diagnosis system for series compensated transmission line based on wavelet transform and adaptive neuro-fuzzy inference system. *Measurement* **2013**, 46, 393-401.
48. Jain, B.; Jain, S.; Nema, R. Investigations on power quality disturbances using discrete wavelet transform. *Int J Electr Electron Comput Eng* **2013**, 2, 47-53.
49. Patil, M.S.M.; Kasturiwala, S.B.; Dahad, S.O.; Jadhav, M.C. Daubechies wavelet tool: Application for human face recognition. *International Journal of Engineering Science and Technology* **2011**, 1, 2392-2398.
50. Stoica, P.; Moses, R.L. *Spectral analysis of signals*. Pearson Prentice Hall Upper Saddle River, NJ: 2005; Vol. 452.
51. Eshel, G. The yule walker equations for the ar coefficients. *Internet resource* **2003**, 2, 68-73.
52. Zhou, H.; Chen, J.; Dong, G.; Wang, H.; Yuan, H. Bearing fault recognition method based on neighbourhood component analysis and coupled hidden markov model. *Mechanical Systems and Signal Processing* **2016**, 66, 568-581.
53. Yang, W.; Wang, K.; Zuo, W. Neighborhood component feature selection for high-dimensional data. *JCP* **2012**, 7, 161-168.
54. Pal, M.; Mather, P.M. An assessment of the effectiveness of decision tree methods for land cover classification. *Remote sensing of environment* **2003**, 86, 554-565.
55. Breiman, L.; Friedman, J.; Olshen, R.; Stone, C. Classification and regression trees, wadsworth international group, belmont, california, USA, 1984; bp roe et al., boosted decision trees as an alternative to artificial neural networks for particle identification. *Nucl. Instrum. Meth. A* **2005**, 543, 57.
56. Moisen, G. Classification and regression trees. **2008**.
57. Kohavi, R. In *A study of cross-validation and bootstrap for accuracy estimation and model selection*, Ijcai, 1995; Stanford, CA: pp 1137-1145.
58. Duda, R.O.; Hart, P.E.; Stork, D.G. *Pattern classification*. John Wiley & Sons: 2012.
59. Vapnik, V.N. An overview of statistical learning theory. *Neural Networks, IEEE Transactions on* **1999**, 10, 988-999.
60. Milgram, J.; Cheriet, M.; Sabourin, R. In "One against one" or "one against all": Which one is better for handwriting recognition with svms?, Tenth International Workshop on Frontiers in Handwriting Recognition, 2006; Suvisoft.
61. Weinberger, K.Q.; Saul, L.K. Distance metric learning for large margin nearest neighbor classification. *Journal of Machine Learning Research* **2009**, 10, 207-244.
62. Xu, Y.; Zhu, Q.; Fan, Z.; Qiu, M.; Chen, Y.; Liu, H. Coarse to fine k nearest neighbor classifier. *Pattern recognition letters* **2013**, 34, 980-986.
63. Kotsiantis, S.B.; Zaharakis, I.; Pintelas, P. Supervised machine learning: A review of classification techniques. 2007.
64. Ho, T.K. The random subspace method for constructing decision forests. *IEEE transactions on pattern analysis and machine intelligence* **1998**, 20, 832-844.
65. Skurichina, M.; Duin, R.P. Bagging, boosting and the random subspace method for linear classifiers. *Pattern Analysis & Applications* **2002**, 5, 121-135.
66. Yang, Y. An evaluation of statistical approaches to text categorization. *Information retrieval* **1999**, 1, 69-90.

67. Hanley, J.A.; McNeil, B.J. The meaning and use of the area under a receiver operating characteristic (roc) curve. *Radiology* **1982**, *143*, 29-36.
68. Breiman, L.; Friedman, J.; Stone, C.J.; Olshen, R.A. *Classification and regression trees*. CRC press: 1984.
69. Demšar, J. Statistical comparisons of classifiers over multiple data sets. *Journal of Machine learning research* **2006**, *7*, 1-30.
70. Garcia, S.; Herrera, F. An extension on "statistical comparisons of classifiers over multiple data sets" for all pairwise comparisons. *Journal of Machine Learning Research* **2008**, *9*, 2677-2694.
71. Conover, W.J.; Iman, R.L. Rank transformations as a bridge between parametric and nonparametric statistics. *The American Statistician* **1981**, *35*, 124-129.
72. Blair, C.R.; Hochberg, Y. Improved bonferroni procedures for testing overall and pairwise homogeneity hypotheses. *Journal of Statistical Computation and Simulation* **1995**, *51*, 281-289.
73. Holm, S. A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics* **1979**, 65-70.
74. Rouder, J.N.; Speckman, P.L.; Sun, D.; Morey, R.D.; Iverson, G. Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic bulletin & review* **2009**, *16*, 225-237.