

Helicopter Main Rotor Fault Diagnosis by Using GA- and PSO- based Classifiers

Soukaina MJAHEd*, Salah EL HADAJ, Khadija BOUZAACHANE, Said RAGHAY

Faculty of Sciences and Technology, Department of Applied Mathematics and Computer Sciences, Cadi Ayyad University, Marrakech, Morocco

mjahedsoukaina1@gmail.com (*Corresponding author)

Abstract: This paper presents an improvement in the recognition of faulty signals, encountered in the case of the Gazelle helicopter's main rotor, using Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) methods. The main focus is on the distinction between faulty and healthy signals and then between the three subclasses of faulty signals, i.e. faulty bearings, joints problem and mechanical loosening. This research work is divided into three parts. The first part approaches the two above-mentioned classes of signals at the same time, and, to this purpose, the Linear Discriminant Analysis (LDA), Non Linear Discriminant Analysis (NLDA) and Back-propagation Neural Network (BPNN) are used. In the second and third part of the paper, GA and PSO are employed for optimizing the hyperplanes and hypersurfaces which separate the above-mentioned classes of signals, as well as the architecture and connection weights of a neural network (NN). Real data are used, which correspond to the vibration signals measured during periodic technical inspections, and are characterized by amplitudes and frequencies typical of the eight highest peaks of the Welch spectrum. The results obtained confirm the validity of the above-mentioned approaches and comparable favorably with those of other multivariate methods. The GA- or PSO-based neural networks' diagnosis can therefore be established for helicopter computers so that faults can be detected.

Keywords: Classification, Genetic algorithm, Particle swarm optimization, Discriminant analysis, Neural networks, Fault diagnosis.

1. Introduction

Anomalous and faulty behaviors damage the machines like engines, gearboxes and rotor systems, causing security reduction, economic losses and environmental damages because of an increased cost of the maintenance and processing time. The fault diagnosis is important for anticipating the incipient fault and predicting system failure in order to ensure the normal operation of an engine.

It is unfeasible that human operators diagnose anomalous faults without this causing them to take wrong decisions, no matter if this happens in a timely manner (Li-Juan et al., 2013; Zhang et al. 2013).

Several researches have introduced the signal recognition and fault diagnosis algorithms. Statistical techniques such as Linear Discriminant Analysis (LDA) (Zhao et al., 2014; Mjahed & Proriol, 1989) were the first to be exploited in this domain. Recently, artificial intelligent techniques, such as Genetic Algorithms (GA) (Mor & Gupta, 2014; Mjahed, 2006), Particle Swarm Optimization (PSO) (Rini et al., 2011; Jena et al., 2015), Artificial Bee Colony Algorithm (Chen & Xiao, 2019), Fuzzy Logic (Raj & Murali, 2013; Xiao et al., 2013) and Artificial Neural Networks (Chandra et al., 2013; Devi & Kumar, 2014) have

been applied successfully to automatic detection and to diagnosis. GA and PSO algorithms have been effectively used to select the attributes of interest (Karimova et al., 2004) and for pattern detection and classification tasks (Hewahi, 2017; Mjahed, 2010).

Characterizing signals with appropriate features and their classification are the most significant issues for machine fault diagnosis (Londhe et al., 2014; Maaref et al., 2018). To identify the faults and achieve a better classification performance, it is important that the features selected contain necessary discriminative information. A number of vibration intensity techniques have been used to analyze engine front noise (Chomphan & Kingrattanaset, 2014). In (Lajmi et al., 2017) a fault diagnosis based on fuzzy Petri net interval is designed.

As regards the fault diagnosis in the aeronautical field, several works have been elaborated. The features for fault diagnosis and prognosis of gearbox are proposed (Zhang et al., 2013). A vibroacoustic technique for the fault diagnosis in a gear transmission of a military helicopter is presented by (Zieja et al., 2017). A genetic algorithm-based neural network for diagnosing an aircraft air compressor bearing is introduced by

(Ouahine et al., 2018). In (Ouahine et al., 2019), a diagnostic system for PUMA helicopter gearbox vibration faults is designed.

In this work, the main focus is on the classification of some specific data corresponding to four classes of faulty signals generated by Gazelle SA341 helicopter's main rotor. In addition to healthy signals, three fault subclasses are considered: the bearing failure, the problems with joints and mechanical loosening. Indeed, these three types of defects remain the most widespread ones (Ouahine et al., 2018, Ouahine et al., 2019; Bhowmik et al., 2013). The considered vibration signals are characterized by 16 discriminant features.

This research work is divided into three parts. The first part is dedicated to the use of LDA, NLDA and BPNN methods. In a second part, the fault signal diagnosis is improved by refining the discriminant functions as well as a neural network through GA technique. In the third part, one attempts to replace GA by PSO for the same goals.

For addressing the topics discussed above, this paper is organized as follows. Section 2 presents the basic definitions of data and signal features. Section 3 is devoted to the implementation of the above-mentioned approaches, where the results of classifications obtained with LDA-, NLDA-, BPNN-, GA/ PSO-based hyperplanes and hypersurfaces, in addition to the GA- and PSO-based neural networks, are detailed. Finally, Section 4 sets forth the main conclusion of this paper.

2. Data and Features Extraction

The data used in this survey have been collected from a Gazelle SA341 helicopter maintenance center and recorded on a test bench. Accelerometers are used to measure the vibrations created by this helicopter's main rotors. Measurements are made during inspection phase.

The faulty signals (denoted as C_1 class) and healthy signals (the C_2 class) are considered. After identifying the fault/healthy signals, the aim is to discriminate between 3 types of faults. The faulty class C_1 is composed of 3 subclasses C_{1i}

(corresponding to problems of the joints), C_{12} (for faulty bearings) and C_{13} (for mechanical loosening).

The collected data amounted to 1000 signals, with 500 learning signals and 500 test signals. Each class participates with 125 learning signals and 125 test signals. All the signals are recorded with a sampling frequency equal to 12000 Hz. Each signal has been divided into 3 segments in order to increase the number of elements in each class. The number of available signals passes over the initial 1000 signals to 3000 signals (375 learning signals and 375 test signals for each class). One explores the Welch estimation method, which provides a consistent estimator of power spectral density (Rani Gupta et al., 2013). This method consists in subdividing data, taking modified periodograms for each section and combining those modified periodograms. In many situations, this method requires less computation than other approaches because it includes converting the signal into shorter sequences.

The spectrum estimate therefore involves estimating the Power Spectral Density based on a finite number of noisy data. Welch's method reduces noise at the expense of frequency resolution (Polat & Güneş, 2007). The Hamming window function has been used in the present case, with a period of 100 and an overlap of 30.

After analyzing the 3000 available signals, the 8 most predominant peaks of the power spectral density have been extracted. For each peak, two values have been used: the characteristic frequency and the spectral density value, which have provided 16 characteristics for each signal. Each signal was modeled with 16 features (v_1 to v_{16}), v_i and v_{i+1} ($i=1, 3, 5, 7, 9, 11, 13, 15$) are the frequency and the spectral density value, respectively, considered in decreasing amplitude values. v_1 and v_2 are the frequency and amplitude of the largest peak, while v_{15} and v_{16} correspond to the eighth peak. The data has then been organized in a matrix of 16 columns and 3000 rows.

3. Fault Diagnosis Approaches

3.1 Classification Methodology

The classification task focuses on the Gazelle SA341 helicopter main rotor fault diagnosis by

considering three faulty signals (C_{11} , C_{12} , C_{13} classes) and a healthy signal (C_2 class).

The same data has amounted to 1500 learning signals, where each class (C_{11} , C_{12} , C_{13} and C_2) has participated with 375 signals and has been used to train the linear and nonlinear discriminant functions as well as the chosen neural networks. Validation data consisting in 1500 other signals, with the same repartition, was used for validating the different methods. It should be noted that, as parameters of performance, the efficiency γ_i and the error ε_i of classifications are computed based N_{24} on the confusion matrix $N=(N_{ij})$, N_{ij} being the value of signals of genuine class C_i classified as class C_j (C_i and C_j mean here classes C_{11} , C_{12} , C_{13} or C_2) (Table 1).

Table 1. Confusion matrix

| Actual class | Predicted class | | | |
|---------------|-----------------|----------|----------|----------|
| | C_{11} | C_{12} | C_{13} | C_2 |
| $C_{11}: N_1$ | N_{11} | N_{12} | N_{13} | N_{14} |
| $C_{12}: N_2$ | N_{21} | N_{22} | N_{23} | N_{24} |
| $C_{13}: N_3$ | N_{31} | N_{32} | N_{33} | N_{34} |
| $C_2: N_4$ | N_{41} | N_{42} | N_{43} | N_{44} |

For each class C_i we have:

$$\begin{cases} \gamma_i = \frac{N_{ii}}{N_i} \\ \varepsilon_i = 1 - \gamma_i \end{cases} \quad (1)$$

The global performance parameters are then calculated as given by (2).

$$\begin{cases} \gamma = \frac{\sum_{i=1}^4 N_{ii} \gamma_i}{\sum_{i=1}^4 N_i} \\ \varepsilon = 1 - \gamma \end{cases} \quad (2)$$

3.2 Part One: LDA-, NLDA- and NN-based Fault Diagnosis

This subsection introduces the achieved LDA, NLDA and BPNN-based fault diagnosis.

3.2.1 LDA-based Diagnosis

3.2.1.1 Basics of LDA

Linear discriminant functions are computed by applying a linear discriminant analysis (LDA) (Mjahed, 2005; Peter He et al., 2005). In the

simple case of 2 classes, with centroids g^1 and g^2 and a covariance matrix M , the expression of the unique discriminant function (or hyperplane) F_l , for a signal vector v , is as given by (3).

$$F_l(v) = (g^1 - g^2)^T M^{-1} v = a^T v \quad (3)$$

This discriminant function is calculated for each pair of classes, through a training sample of vibration signals. This is the purpose of the following subsection.

3.2.1.2 LDA-based Diagnosis Implementation

Applied to this learning data set, the best discriminant functions have been sought according to (3), for best efficiency.

Then three best linear discriminant functions are obtained (4): F_l (for the discrimination between C_1 and C_2 signals), F_{11} (for the two faulty classes C_{11} and $C_{12} \cup C_{13}$) and F_{12} (to classify C_{12} based on C_{13} faulty signals).

The best discriminant functions correspond to the most efficient functions. In fact, as illustrated by (4), the most efficient functions are those based on the 6 variables $v_1 - v_6$, i.e. the coefficients corresponding to the variables $v_7 - v_{16}$ are almost zero.

$$\begin{cases} F_l(v) = -1.12 + 1.3v_1 + 1.8v_2 + 1.6v_3 + 2.1v_4 \\ \quad + 1.5v_5 + 2.1v_6 \\ F_{11}(v) = -1.91 + 3.1v_1 + 1.1v_2 + 0.9v_3 + 3.4v_4 \\ \quad + 2.3v_5 + 1.9v_6 \\ F_{12}(v) = -2.23 + 1.1v_1 + 0.8v_2 + 0.6v_3 + 0.1v_4 \\ \quad + 2.6v_5 + 1.2v_6 \end{cases} \quad (4)$$

3.2.2 NLDA-based Diagnosis

3.2.2.1 Basics of NLDA

In order to improve the separation between the considered classes C_i and C_j , LDA is generalized to NLDA, by going from the first degree (3) to the second degree as given in (5) (Kuhn & Johnson, 2016):

$$\begin{cases} F(v) = (a^T + v^T B)v \\ a^T = (g^1 - g^2)^T M^{-1} \\ B = (G^1 - G^2)^T H^{-1} \end{cases} \quad (5)$$

By expanding a to contain the elements of B and by expanding v to contain the elements of vv^T , the linear deduction of the coefficients may be used. G^1 , G^2 and H are the new centroids for classes (C_1 and C_2) and the covariance matrix for the modified data, respectively.

3.2.2.2 NLDA-based Diagnosis Implementation

Here a quadratic form has been chosen, which consists in calculating parameter B (5) for each combination of attributes. The optimal results are given in (6), where F_2 , F_{21} and F_{22} are hypersurfaces separating the C_1 and C_2 signals, C_{11} and $C_{12} \cup C_{13}$ signals and the C_{12} and C_{13} classes, respectively. As mentioned above, the best discriminant functions are those combining the variables v_1 - v_6 .

$$\begin{cases} F_2(v) = -0.23 + 1.1v_1^2 + 0.8v_2^2 + 0.6v_3^2 + 0.1v_4^2 \\ \quad + 0.8v_5^2 + 0.3v_6^2 \\ F_{21}(v) = -0.61 + 2.5v_1^2 + 1.2v_2^2 + 2.4v_3^2 + 1.3v_4^2 \\ \quad + 1.1v_5^2 + 0.7v_6^2 \\ F_{22}(v) = -1.24 + 0.9v_1^2 + 2.5v_2^2 + 1.7v_3^2 + 2.1v_4^2 \\ \quad + 0.9v_5^2 + 0.6v_6^2 \end{cases} \quad (6)$$

3.2.3 BPNN-based Fault Diagnosis

3.2.3.1 BPNN Principle

The architecture of a multilayered neural network is organized into levels of neurons: one input layer, one output layer and one or several hidden layers (Haykin, 2009). Each neuron i in a level l is thus directly connected to all the neurons of the following layer ($l+1$) and produces a response $Y_i^{(l)}$ after being processed through a nonlinear sigmoid function g (7).

$$Y_i^{(l)} = g\left(\sum_{j=1}^{N_{l-1}} W_{ij}^{(l)} Y_j^{(l-1)} - \theta_i^{(l)}\right) \quad (7)$$

$\theta_i^{(l)}$ is the threshold of the neuron i in layer l , $W_{ij}^{(l)}$ the connection weight between the neurons i (of layer l) and j (of layer $l-1$), N_{l-1} the number of nodes in the layer $l-1$ and $Y_j^{(l-1)}$ the neuron j of layer ($l-1$) output.

Generally, the training of this network is supervised and uses an error back-propagation (BP) algorithm. The network seeks to minimize, on the output layer, and for each signal data p , a

quadratic error $E^{(p)}$ (8) which exists between the effective output value $o_i^{(p)}$ and the desired output value $d_i^{(p)}$ of the m neurons.

$$E = \sum_p E^{(p)} = \sum_p \sum_{i=1}^m (o_i^{(p)} - d_i^{(p)})^2 \quad (8)$$

3.2.3.2 BPNN-based Diagnosis Implementation

As it has been introduced above, the identification of Helicopter main rotor fault signals belonging to C_{11} , C_{12} , C_{13} and C_2 classes has been accomplished by means of “3-layer” neural network. The optimized architecture is reflected by (6, 10, 2).

The number of neurons in the input layer which is set at 6 is linked to the number of variables reduced from 16 to 6. The tests on the number of neurons in the hidden layer have shown that 10 neurons are sufficient to converge the neural network faster and with a very small error. In this paper, two neurons are used for the output layer. Each signal p described based on the above defined variables, $X = (v_1, v_2, v_3, v_4, v_5, v_6)$, was processed through the input layer, which in turn fed it to the hidden layer, thereby feeding it finally to two output neurons o_1 and o_2 .

The 80 synaptic weights W_{ij} between neurons and 10 thresholds θ_i are computed. These parameters are adjusted in such a way that the error function E (8) is minimized using the back-propagation algorithm. The desired outputs (d_1, d_2) for the four classes C_{11} , C_{12} , C_{13} and C_2 are $(+1, +1)$, $(+1, -1)$, $(-1, +1)$ and $(-1, -1)$, respectively.

3.3 Part Two: GA-based Fault Diagnosis

3.3.1 GA Principle

The genetic algorithms are based on the principles of genetics and Darwin’s natural selection theory (“the one that is best endowed, survives”). Simple GA is based on three essential operators: selection, crossover and mutation (Mor & Gupta, 2014; Mjahed, 2010).

GA operates on a population consisting of a number of prospective solutions. The population at time t is characterized by the time-dependent variable $S(t)$, with the random initial population $S(0)$. The design parameters to be searched are

the real components of $S(t)$. $S(t) = (S_{ij}(t)); (i=\{1, \dots, N_{ind}\}, j=\{1, \dots, N_{par}\})$, N_{ind} being the population size and N_{par} the number of parameters to be optimized.

Based on the specified parameter ranges, these components are then developed generation by generation and new solutions are gained. After a number of generations (N_{gen}), the GA is completed. The best solution corresponds to the minimal fitness function.

3.3.2 Fault Diagnosis Using GA-based Hyperplanes (GAHP)

According (9) it can be proposed to seek the three GA-based hyperplanes (GAHP), namely F_3 , F_{31} and F_{32} , separating between the classes (C_1 and C_2), (C_{11} and $C_{12} \cup C_{13}$) and (C_{12} and C_{13}) respectively.

$$\begin{cases} F_3(v) = \alpha_0 + \sum_{i=1}^{16} \alpha_i v_i \\ F_{31}(v) = \alpha_{01} + \sum_{i=1}^{16} \alpha_{i1} v_i \\ F_{32}(v) = \alpha_{02} + \sum_{i=1}^{16} \alpha_{i2} v_i \end{cases} \quad (9)$$

where v_i ($i=\{1, \dots, 16\}$) are the variables defined above.

For this approach, the global misclassification rate ε has been used as fitness function.

The parameters to be explored, by using GA, are the 51 (3×17) factors of F_3 , F_{31} and F_{32} ($\alpha_i, \alpha_{i1}, \alpha_{i2}$ ($i=\{0, \dots, 16\}$)). To accomplish this search, one uses a population of 20 solutions and 1000 generations, a crossover probability equal to 0.8, a mutation probability equal to 0.3, a mutation rate equal to 0.02 and a selection pressure equal to 8. The best GAHP-based discriminant functions (best results over 100 runs) are given in (10). The results obtained once again unveil functions with 6 variables, illustrating the predominance of variables v_1-v_6 . Here too, the coefficients relating to the variables v_7-v_{16} are practically zero.

$$\begin{cases} F_3(v) = -2.01 + 1.1v_1 + 1.9v_2 + 0.6v_3 + 1.2v_4 \\ \quad + 1.3v_5 + 2.3v_6 \\ F_{31}(v) = -3.15 + 1.2v_1 + 1.8v_2 + 1.1v_3 + 0.8v_4 \\ \quad + 1.9v_5 + 1.5v_6 \\ F_{32}(v) = -2.63 + 1.4v_1 + 1.4v_2 + 0.9v_3 + 1.4v_4 \\ \quad + 2.1v_5 + 1.9v_6 \end{cases} \quad (10)$$

3.3.3 Fault Diagnosis by GA-based Hypersurfaces (GAHS)

After some investigations, the simple expression given by (11) was selected, which has the same form as (6). F_4 , F_{41} and F_{42} are three hypersurfaces separating between classes (C_1 and C_2), (C_{11} and $C_{12} \cup C_{13}$) and (C_{12} and C_{13}) respectively.

$$\begin{cases} F_4(v) = \rho_0 + \sum_{i=1}^{16} \rho_i v_i^2 \\ F_{41}(v) = \rho_{01} + \sum_{i=1}^{16} \rho_{i1} v_i^2 \\ F_{42}(v) = \rho_{02} + \sum_{i=1}^{16} \rho_{i2} v_i^2 \end{cases} \quad (11)$$

For the three surfaces mentioned above, the parameters to be searched are the 51 aforementioned coefficients ($\rho_i, \rho_{i1}, \rho_{i2}$ ($i=\{0, \dots, 16\}$)). The same parameters expressed in GAHP search have been used for GAHS. The best GAHS-based discriminant functions obtained over 100 runs are expressed by (12), revealing again the relevance of the variables v_1-v_6 .

$$\begin{cases} F_4(v) = -1.93 + 1.4v_1^2 + 2.1v_2^2 + 0.8v_3^2 + 1.5v_4^2 \\ \quad + 1.6v_5^2 + 0.9v_6^2 \\ F_{41}(v) = -4.11 + 0.8v_1^2 + 2.6v_2^2 + 1.2v_3^2 + 1.1v_4^2 \\ \quad + 0.9v_5^2 + 1.1v_6^2 \\ F_{42}(v) = -3.74 + 0.7v_1^2 + 1.5v_2^2 + 1.4v_3^2 + 2.3v_4^2 \\ \quad + 2.1v_5^2 + 1.6v_6^2 \end{cases} \quad (12)$$

3.3.4 Fault Diagnosis by GA-based Neural Network (GANN)

In subsection 3.2.3.2, a neural network has been optimized to achieve the architecture reflected by (6, 10, 2). Here, for a NN architecture ($n_p, n_h, 2$), GA is used to optimize the neuron numbers (n_p, n_h) in the input and hidden layers as well as the different connection weights and thresholds.

The number of parameters to be optimized is therefore $n_p n_h + 2n_h$ connection weights, in addition to n_h thresholds. To carry out this task, one uses a population of 50 solutions, 1000 generations, a crossover probability equal to 0.8, a mutation probability equal to 0.3, a mutation rate equal to 0.02 and a selection pressure equal to 8. The GANN error E is used as a fitness function. The best architecture obtained is identical to that obtained in the previous section as reflected by (6, 10, 2).

3.4 Part Three: PSO-based Fault Diagnosis

3.4.1 PSO Basic Notions

Particle swarm optimization (PSO) is the most well-known and popular nature-inspired optimization algorithm.

By means of the PSO algorithm, a swarm of particles (solutions) is randomly initialized (position (x) and velocity (ϕ)). The movement of the particles is directed by their own best-known position x_{pbest} and the swarm best-known position x_{gbest} which are related to the best solution (fitness) (Rini et al., 2011; Jena et al., 2015). For every generation, and at each time step t , the position and the velocity of each particle are updated according to (13)-(14):

$$x(t) = x(t-1) + \phi(t) \quad (13)$$

$$\begin{aligned} \phi(t) = & w_t \phi(t-1) + c_1 R_1 (x_{pbest} - x(t-1)) \\ & + c_2 R_2 (x_{gbest} - x(t-1)) \end{aligned} \quad (14)$$

where c_1 and c_2 are the constants designated as acceleration/learning factors that represent the weighing of the stochastic terms pulling each particle toward the x_{pbest} and x_{gbest} positions respectively. R_1 and R_2 are obtained by a uniform distribution in the interval [0 1]. The inertia weight w_t is calculated, using (15).

$$w_t = 0.05 w_{t-1} w_{damp} \quad (15)$$

where w_{damp} is the inertia weight damping ratio.

3.4.2 Fault Diagnosis by PSO-based Hyperplanes (PSOHP)

In this subsection, PSO is employed for determining three hyperplanes as defined by (16). These hyperplanes F_5 , F_{51} and F_{52} are separating the classes (C_1 and C_2), (C_{11} and $C_{12} \cup C_{13}$) and (C_{12} and C_{13}) respectively.

$$\begin{cases} F_5(v) = \delta_0 + \sum_{i=1}^{16} \delta_i v_i \\ F_{51}(v) = \delta_{01} + \sum_{i=1}^{16} \delta_{i1} v_i \\ F_{52}(v) = \delta_{02} + \sum_{i=1}^{16} \delta_{i2} v_i \end{cases} \quad (16)$$

where v_i ($i=1, \dots, 16$) are the 16 features already defined.

The parameters to be optimized, by PSO, are then the 51 aforementioned coefficients ($\delta_p, \delta_{ip}, \delta_{i2}$ ($i=1, \dots, 16$)). The PSO parameters have been efficiently tuned in order to get optimal results. This method is run up to 1000 iterations and uses a population size of 20 solutions. In this approach, the learning factors c_1 and c_2 , which are both equal to 1.49, have been used. The linearly decreasing inertia weight (w_t) was implemented using an inertia weight damping ratio $w_{damp} = 0.99$.

The best PSOHP-based discriminant functions obtained are given by (17), also showing the significance of the first 6 variables. The coefficients corresponding to v_7-v_{16} in (16) are zero.

$$\begin{cases} F_5(v) = -2.62 + 1.2v_1 + 0.8v_2 + 1.2v_3 + 1.3v_4 \\ \quad + 1.7v_5 + 2.4v_6 \\ F_{51}(v) = -3.41 + 1.8v_1 + 1.2v_2 + 0.9v_3 + 1.2v_4 \\ \quad + 1.9v_5 + 1.1v_6 \\ F_{52}(v) = -4.13 + 1.7v_1 + 1.6v_2 + 1.6v_3 + 1.1v_4 \\ \quad + 2.1v_5 + 1.7v_6 \end{cases} \quad (17)$$

3.4.3 Fault Diagnosis by PSO-based Hypersurfaces (PSOHS)

To separate between the two classes, one uses the same expression utilised for the NLDA and GAHS approaches, as given in (18), where F_6 , F_{61} and F_{62} are three hypersurfaces separating classes (C_1 and C_2), (C_{11} and $C_{12} \cup C_{13}$) and (C_{12} and C_{13}) respectively.

$$\begin{cases} F_6(v) = \sigma_0 + \sum_{i=1}^{16} \sigma_i v_i^2 \\ F_{61}(v) = \sigma_{01} + \sum_{i=1}^{16} \sigma_{i1} v_i^2 \\ F_{62}(v) = \sigma_{02} + \sum_{i=1}^{16} \sigma_{i2} v_i^2 \end{cases} \quad (18)$$

This approach aims to optimize the 51 aforementioned coefficients ($\sigma_p, \sigma_{ip}, \sigma_{i2}$ ($i=0, \dots, 16$)) using PSO algorithm tuned with the same parameters defined in the previous subsection. The best PSOHS-based discriminant functions are grouped in (19), allowing to make the same observation as previously concerning the weight of the variables v_1-v_6 .

$$\begin{cases} F_6(v) = -2.23 + 1.5v_1^2 + 1.5v_2^2 + 0.7v_3^2 + 2.1v_4^2 \\ \quad + 1.6v_5^2 + 0.9v_6^2 \\ F_{61}(v) = -3.61 + 2.5v_1^2 + 1.9v_2^2 + 1.1v_3^2 + 1.9v_4^2 \\ \quad + 2.4v_5^2 + 1.1v_6^2 \\ F_{62}(v) = -4.24 + 0.9v_1^2 + 2.1v_2^2 + 0.9v_3^2 + 2.3v_4^2 \\ \quad + 0.9v_5^2 + 1.2v_6^2 \end{cases} \quad (19)$$

3.4.4 Fault Diagnosis by PSO-based Neural Network (PSONN)

It should be mentioned that for the PSO-based neural network (PSONN) approach, the same analysis carried out for GANN has been extended. PSO is employed for optimizing the $(n_p, n_h, 2)$ NN architecture and the connection weights and thresholds. The PSONN method is run up to 1000 iterations and it uses a population size of 50 solutions, with the same number of optimization parameters as the GANN approach. The learning factors c_1 and c_2 are set to be equal to 1.49. The inertia weight damping ratio is 0.99.

Here again, the best architecture is: (6, 10, 2).

3.5 Decision Rules

It should be noted that, after calculating or optimizing functions F_i, F_{i1} and F_{i2} , ($i=1, \dots, 6$) based on training set, in the framework of the LDA, NLDA, GAHP, GAHS, PSOHP and PSOHS approaches, a test signal v_0 is assigned to one of the four classes according to decision rule (20).

$$\left\{ \begin{array}{l} \text{if } F_i(v_0) \geq 0 \text{ then} \\ \quad \text{if } F_{i1}(v_0) \geq 0 \text{ then } v_0 \in C_{11} \\ \quad \text{elseif } F_{i2}(v_0) \geq 0 \text{ then } v_0 \in C_{12} \\ \quad \text{else } v_0 \in C_{13} \\ \quad \text{endif} \\ \text{else } v_0 \in C_2 \\ \text{endif} \end{array} \right. \quad (20)$$

It should be mentioned that in order to conduct a proper comparison, the same form of discriminant functions ((4), (10), (17)) have been used for LDA, GAHP, and PSOHP approaches and another same form of non-linear discriminant functions ((6), (12), (19)) have been selected for NLDA, GAHS and PSOHS approaches.

For establishing a diagnosis linked to the use of neural networks (BPNN, GANN and PSONN approaches), and since the desired neural networks' outputs (d_1, d_2) for the four classes C_{11}, C_{12}, C_{13} and C_2 are $(+1, +1), (+1, -1), (-1, +1)$ and $(-1, -1)$, respectively, this paper proposes a decision rule for every faulty signal v_0 , depending on effective NN outputs as indicated in (21).

$$\left\{ \begin{array}{l} \text{if } o_1(v_0) \geq 0 \text{ and } o_2(v_0) \geq 0 \text{ then } v_0 \in C_{11} \\ \text{if } o_1(v_0) \geq 0 \text{ and } o_2(v_0) < 0 \text{ then } v_0 \in C_{12} \\ \text{if } o_1(v_0) < 0 \text{ and } o_2(v_0) \geq 0 \text{ then } v_0 \in C_{13} \\ \text{if } o_1(v_0) < 0 \text{ and } o_2(v_0) < 0 \text{ then } v_0 \in C_2 \end{array} \right. \quad (21)$$

3.6 Validation Results and Analysis

This section is devoted to the validation of the discriminant functions given by (4), (6), (10), (12), (17) and (19) as well as the optimal BPNN, GANN, PSONN neural networks, with the (6, 10, 2) architecture, obtained in subsections 3.2.3.2, 3.3.4 and 3.4.4 respectively.

The validation results for the best discriminant functions and neural networks, using the validation data, are achieved by means of classification matrices, based on which the efficiency rates are extracted and reported in Table 2.

Since PSO and GA are stochastic optimization methods, they produce different results in different runs. In this sense, the discriminant functions (10), (12), (17) and (19) illustrate the best results obtained over 100 runs based on the GAHP, GAHS, PSOHP and PSOHS approaches.

Table 2. Efficiency Rates γ_i obtained by using LDA, GAHP, PSOHP, NLDA, GAHS, PSOHS, BPNN, GANN and PSONN approaches (validation data)

| Method | γ_i (%) | | | | |
|--------|----------------|--------------|--------------|--------------|--------------|
| | C_{11} | C_{12} | C_{13} | C_2 | All |
| LDA | 91.73 | 92.80 | 91.46 | 92.26 | 92.06 |
| GAHP | 94.13 | 94.66 | 95.46 | 95.20 | 94.86 |
| PSOHP | 93.86 | 94.93 | 94.66 | 95.46 | 94.72 |
| NLDA | 97.06 | 97.60 | 97.06 | 96.00 | 96.93 |
| GAHS | 98.13 | 98.93 | 97.60 | 97.60 | 98.06 |
| PSOHS | 98.13 | 98.40 | 97.60 | 97.60 | 97.93 |
| BPNN | 98.93 | 99.20 | 98.66 | 98.66 | 98.86 |
| GANN | 99.20 | 99.46 | 99.20 | 99.46 | 99.33 |
| PSONN | 98.93 | 99.20 | 98.93 | 99.20 | 99.06 |

The first observation concerns the number of attributes required for a good classification of the different faults. When computing or optimizing discriminant functions by GA and PSO, it turned out that six attributes ($v_1, v_2, v_3, v_4, v_5, v_6$) were sufficient to achieve the best results (formulas (4),(6), (10), (12), (17) and (19)). In fact, the six best variables are those corresponding to the first

3 peaks of the power spectral density. This finding has also been validated by forcing combinations with more than six attributes (see Figures 1a-c).

The Figures 1(a)-(c) illustrate the evolution of the classifications' efficiency rates according to the number of attributes involved.

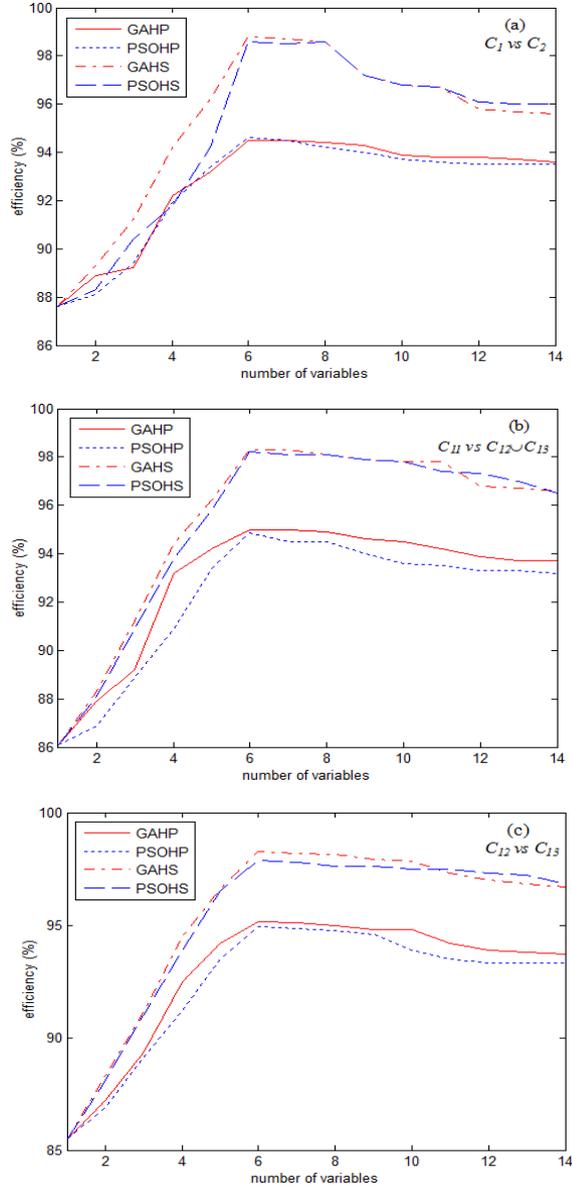


Figure 1. Efficiency vs number of variables in classification (a) between healthy (C_1) and faulty (C_2) signals, (b) between faulty signals ($C_{11} / C_{12} \cup C_{13}$) and (c) between (C_{12} / C_{13}), obtained based on GAHP, PSOHP, GAHS and PSOHS approaches

For the separation between healthy main rotors (C_1) and faulty main rotors (C_2), the efficiency rates do not increase beyond 6 attributes, for GAHP, PSOHP, GAHS, PSOHS approaches (Figure 1a). The same can be noticed with regard to the classification between three faulty behaviors ($C_{11} / C_{12} \cup C_{13}$) in Figure 1b and (C_{12} / C_{13}) in

Figure 1c. In fact, efficiency rates tend to increase as the number of variables used increases. Beyond six variables, these rates become almost constant and tend to decrease in some cases.

Even in the case of the neural approach, the input layer is more efficient when it contains six neurons.

With respect to BPNN optimization, Figure 2a shows the evolution of the BPNN efficiency with architecture ($n, 10, 2$) as a function of the number n of neurons in input layer (for n ranging from 4 to 14). Similarly, an analysis was performed by plotting the BPNN efficiency variation with respect to n (the hidden layer neuron number), with n taking values between 4 and 20.

Figure 2b illustrates the decrease of the (6, 10, 2) BPNN error as a function of epochs. This neural network converged after 25 epochs with a mean squared error $E = 0.00006$.

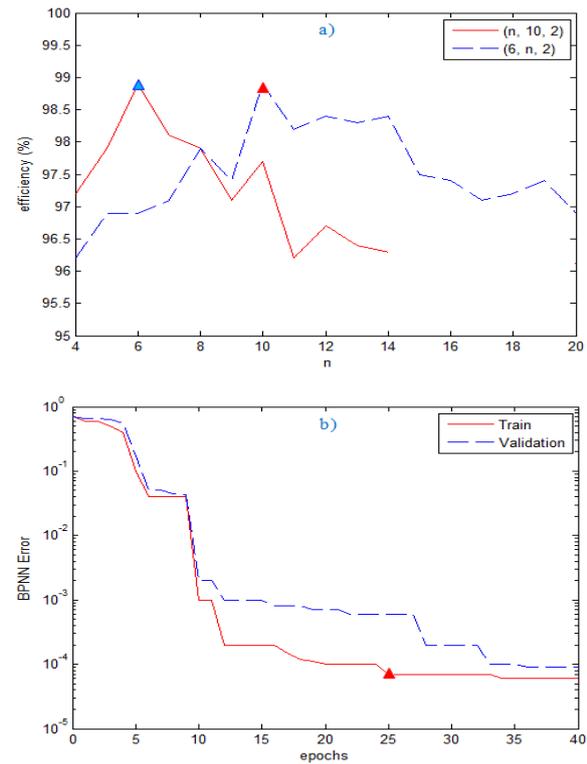


Figure 2. (a) BPNN efficiency vs n for ($n, 10, 2$) and (6, $n, 2$) architectures, (b) BPNN training and validation error for (6, 10, 2) architecture

As suggested by the GANN and PSANN approaches, a fitness evolution as a function of n_1 and n_2 (the number of neurons in input and hidden layers) is shown in Figures 3-4, where n_1 ranges from 1 to 14, and n_2 ranges from 1 to 25.

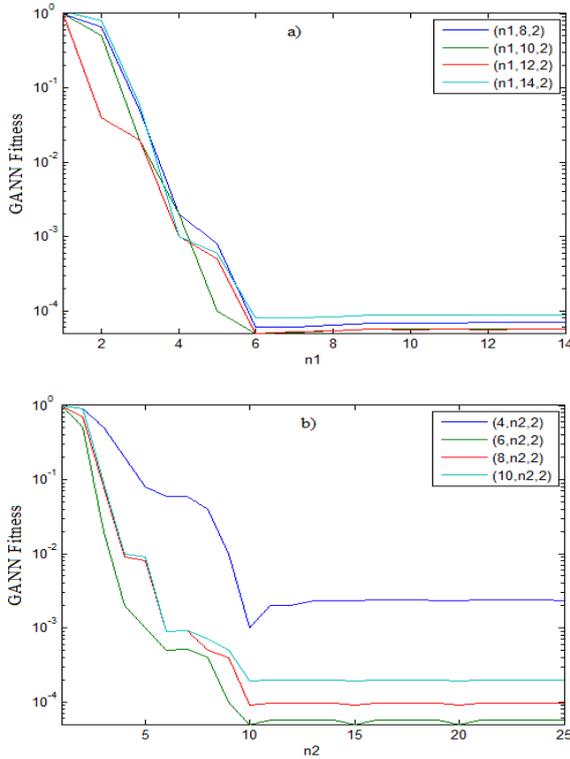


Figure 3. GANN fitness as a function of n_1 (a) and n_2 (b) for different NN architectures

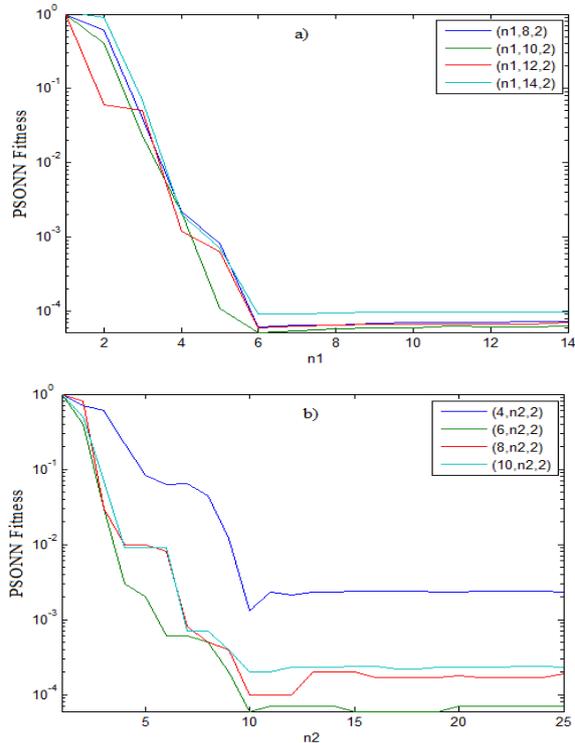


Figure 4. (a) PSONN fitness as a function of n_1 and (b) PSONN fitness as a function of n_2 (for different NN architectures)

The illustrated results express the accuracy of the optimal architecture (6, 10, 2) regardless of the approach used. It should be noted that the

efficiency and fitness plots as functions of the number of variables (Figures 1, 2a, 3a and 4a) are given up to 14 variables.

The fault diagnosis performed by neural networks is more efficient compared to other methods. In this case, it would be wise to mention a slight improvement for GANN and PSONN compared to BPNN.

The efficiency of the classifications obtained via the GA and the PSO algorithms are clearly better. The efficiencies of the GA and PSO algorithms are similar. Moreover, the classifications based on hypersurfaces (NLDA, GAHS, PSOHS) are more efficient than the ones based on hyperplanes (LDA, GAHP, PSOHP).

The computation code for these different analyzes was compiled under the Matlab environment (R2017b) on an Intel Core i73.0 GHz processor, with 8GB of RAM. The execution times of the considered approaches, carried out offline during the training step, are mentioned in Table 3.

Concerning the choice of attributes, the present results indicate the relevance of the amplitudes and frequencies of the spectrum peaks compared to other types of attributes generally used (Li-Juan et al. 2013; Zhang et al. 2013). Given the complexity of the classification task, these results compare favorably to those obtained in other works (Raj & Murali, 2013; Xiao et al., 2013). Due to the employment of neural networks and their GA- and PSO-based hybridization used in the aeronautical fault diagnosis, results comparable to those published in the works (Londhe et al., 2014; Chomphan & Kingrattanaset, 2014) have been obtained.

Table 3. Run times for LDA, GAHP, PSOHP, NLDA, GAHS, PSOHS, BPNN, GANN and PSONN Approaches (training step)

| Method | LDA | GAHP | PSOHP |
|----------------|------|------|-------|
| CPU Time (min) | 1 | 3 | 3 |
| Method | NLDA | GAHS | PSOHS |
| CPU Time (min) | 3 | 7 | 7 |
| Method | BPNN | GANN | PSONN |
| CPU Time (min) | 8 | 14 | 14 |

4. Conclusion

The aim of this paper is to use the Genetic Algorithms and Particle Swarm Optimization methods to improve the discrimination between faulty and healthy signals generated by Gazelle SA341 helicopter's main rotors.

GA and PSO were exploited in order to improve the discriminant functions' parameters as well as the neural network architecture and parameters. A database of vibratory signals collected during periodic inspections has been used for this purpose. It can be easily noticed, that all the results obtained by using the GA- and PSO-based techniques display a higher efficiency and therefore a higher discrimination ability.

The LDA optimization is comparable to the hyperplane search approaches (GAHP and PSOHP) and the NLDA optimization is comparable to the hypersurface search approaches (GAHS and PSOHS). According to Table 2, GAHP and PSOHP outclass LDA with total efficiency rates of 94.86%, 94.72% and 92.06%, respectively. Similarly, GAHS and PSOHS perform better than NLDA with total efficiency rates approaching 98.06%, 97.93% and 96.79%, respectively.

The GA- and PSO-based Neural Networks have shown better results and performances than the discriminant analysis. PSONN outperforms BPNN with efficiency rates of 99.06% and 98.86% (respectively). Very satisfactory results have been obtained by using the GANN approach, with a classification rate of 99.33% for validation data. The GANN method was validated and retained for a possible computer implantation for main rotor faults' diagnosis, considering in this regard more than three fault sub-classes. The overall performances are interesting and in agreement with the complexity of the above-mentioned classification task. Based on all simulation scores, it can be noticed that GA and PSO have been able to find optimal factors of discriminant functions and neural networks. These results are promising and suggest that GA and PSO can also be employed for other complex designs. GA and PSO allow one to improve the efficiency rates of classifications and to minimize classification errors.

GANN (or PSONN) is characterized by quick diagnostic decision-making and ease of implementation on ground-based computers or in-flight systems. Its advantage lies in its ability to know the state of the helicopter main rotor without making it necessary to disassemble its elements.

REFERENCES

- Bhowmik, P. S., Pradhan, S. & Prakash, M. (2013). Fault Diagnostic and Monitoring Methods of Induction Motor: A Review, *International Journal of Applied Control, Electrical and Electronics Engineering*, 1(1), 1-18.
- Chandra, M., Sekhar, R. & Sekhar, A. S. (2013). Application of Artificial Neural Networks for Identification of Unbalance and Looseness in Rotor Bearing Systems, *International Journal of Applied Science and Engineering*, 11(1), 69-84.
- Chen, W. & Xiao, Y. (2019). An Improved ABC Algorithm and its Application in Bearing Fault Diagnosis with EEMD, *Algorithms*, 12(4), 72, 1-12.
- Chomphan, S. & Kingrattanaset, T. (2014). An Analysis of Sound for Fault Engine, *American Journal of Applied Sciences*, 11(6), 1005-1009.
- Devi, S. & Kumar, L. S. (2014). An Innovative Approach in Fault Diagnostics of Steam Turbine Performance using ANN and M-SVM Algorithms, *International Review of Mechanical Engineering*, 8(2), 350-359.
- Haykin, S. (2009). *Neural Networks and Learning Machines*, Third Edition, Pearson Education.
- He, P. Q., Qin, J. S. & Wang, J. (2005). A New Fault Diagnosis Method using Fault Directions in Fisher Discriminant Analysis, *American Institute of Chemical Engineers Journal*, 51(2), 555-571.
- Hewahi, N. (2017). A Hybrid Approach Based on Genetic Algorithm and Particle Swarm Optimization to Improve Neural Network Classification, *Journal of Information Technology Research*, 10(3), 48-68.
- Jena, P. K., Thatoi, D. N. & Parhi, D. R. (2015). Dynamically Self-Adaptive Fuzzy PSO Technique for Smart Diagnosis of Transverse Crack, *Applied Artificial Intelligence*, 29(3), 211-232.

- Karimova, L., Kuadykova, E. & Makarenkoa, N. (2004). The Genetic Algorithm for a Signal Enhancement, *Nuclear Instruments and Methods, A* 534(2004), 170-174.
- Kuhn, M. & Johnson, K. (2016). *Applied Predictive Modeling*, Springer: New York, NY, USA.
- Lajmi, F., Talmoudi, A. J. & Dhouibi, H. (2017). Fault Diagnosis of Uncertain Systems Based on Interval Fuzzy PETRI Net, *Studies in Informatics and Control*, 26(2), 239-248. DOI: 10.24846/v26i2y201712
- Li-Juan, G., Chun-hui, Z., Min, H. & Yong, Z. (2013). Vibration Analysis of the Steam Turbine Shafting caused by Steam Flow, *Telkomnika*, 11(8), 4422-4432.
- Londhe, P. S., Patre, B. M. & Tiwari, A. P. (2014). Design of Single-Input Fuzzy Logic Controller for Spatial Control of Advanced Heavy Water Reactor, *IEEE Transactions on Nuclear Science*, 61(2), 901-911.
- Maaref, B., Ghabi, J. & Simeu-Abazi, Z. (2018). A Mixed Approach for Modeling and Fault Diagnosis of Hybrid Dynamical Systems with Unknown Disturbances, *Studies in Informatics and Control*, 27(3), 307-318. DOI: 10.24846/v27i3y201806
- Mjahed, M. & Proriol, J. (1989). Tagging Jets and Partons by Discriminant Analysis, *Physics Letters B*, 217(4), 560-562.
- Mjahed, M. (2005). Higgs Search at LHC by Neural Networks, *Nuclear Physics B*, 140, 799-801.
- Mjahed, M. (2006). Search for the Higgs boson at LHC by using Genetic Algorithms, *Nuclear Instruments and Methods, A* 559(1), 172-178.
- Mjahed, M. (2010). Optimization of Classification Tasks by using Genetic Algorithms. In *Proceedings of Informatics and Systems (INFOS)*, IEEE Conference, Cairo, Egypt, March 28-30 (pp. 1-4).
- Mor, M. & Gupta, P. (2014). A Review on Clustering with Genetic Algorithms, *International Journal of Computer Science & Communications Networks*, 4(3), 94-98.
- Ouadine, A. Y., Mjahed, M., Ayad, H. & El Kari, A. (2018). Aircraft Air Compressor Bearing Diagnosis Using Discriminant Analysis and Cooperative Genetic Algorithm and Neural Network Approaches, *Applied Sciences*, 8(2243), 1-16.
- Ouadine, A. Y., Mjahed, M., Ayad, H. & El Kari, A. (2019). Helicopter gearbox vibration fault classification using order tracking method and genetic algorithm, *Automatika*, 60(1), 68-78.
- Polat, K. & Güneş, S. (2007). Classification of Epileptiform EEG using a Hybrid System based on Decision Tree Classifier and Fast Fourier Transform, *Applied Mathematics and Computation*, 187(2), 1017-1026.
- Raj, A. S. & Murali, N. (2013). Early Classification of Bearing Faults Using Morphological Operators and Fuzzy Inference, *IEEE Transactions on Industrial Electronics*, 60(2), 567-574.
- Rani Gupta, H., Batan, S. & Mehra, R. (2013). Power Spectrum Estimation using Welch Method for various Window Techniques, *International Journal of Scientific Research Engineering & Technology (IJSRET)*, 2(6), 389-392.
- Rini, P., Shamsuddin, M. & Yuhaniz, S. (2011). Particle Swarm Optimization: Technique, System and Challenges, *International Journal of Computer Applications*, 14, 19-27.
- Xiao, H., Zhou, J., Xiao, J. & Fu, W. (2013). Identification of Vibration-Speed Curve for Hydroelectric Generator unit using Statistical Fuzzy Vector Chain Code and SVM machine. In *Proceedings of the Institution of Mechanical Engineers Part O-Journal of Risk and Reliability*, 228(3) (pp. 291-300).
- Zhang, X., Kang, J., Zhao, J. S. & Cao, D. C. (2013). Features for Fault Diagnosis and Prognosis of Gearbox, *Chemical Engineering Transactions*, 33, 1027-1032.
- Zhao, M., Jin, X., Zhang, Z. & Li, B. (2014). Fault Diagnosis of Rolling Element Bearings via Discriminative Subspace Learning: Visualization and Classification, *Expert Systems with Applications*, 41(7), 3391-3401.
- Zieja, M., Golda, P., Zokowski, M. & Majewski, P. (2017). Vibroacoustic Technique for the Fault Diagnosis in a Gear Transmission of a Military Helicopter, *Journal of Vibroengineering*, 19(2), 1039-1049.