Application of optimised neural networks models in gears and bearings faults diagnosis

Kaaïs Khoualdia*
Faculty of Science and Technology,
Department of Mechanical Engineering,
Souk Ahras University,
Souk Ahras 41 000, Algeria
Email: khoualdiaakis@yahoo.com
*Corresponding author

Elias Hadjadj Aoul
Department of Electromechanical,
Electromechanical System Laboratory,
Annaba University,
Annaba 23 000, Algeria
Email: hadjadj.elias@yahoo.fr

Tarek Khoualdia
Faculty of Science and Technology,
Department of Mechanical Engineering,
Souk Ahras University,
Souk Ahras 41 000, Algeria
Email: khoualdiaatarek@hotmail.fr

Abstract: Gears and bearings are some of the most important machine components in the industrial world and detection of their faults has become a major trend. In the present article, in order to bring a reliable methodology for monitoring and diagnosis of rotating machinery failures, a test rig is implemented. However, to diagnose gears and bearings combined faults, a monitoring system based on neural network model (NNM), is proposed. To train and test the NNM, the principal high frequency indicators, determined with the collected time domain vibration data, and codes of defects are used respectively as input and output data. A comparison of two learning algorithms, optimised by the Taguchi method, was done to determine the best NNM. Therefore, the proposed method is effective to study other various industrial cases.

Keywords: gear and bearing combined defects; fault vibration analysis; ANN; artificial neural network; design of experiment; Taguchi method.

Reference to this paper should be made as follows: Khoualdia, K., Aoul, E.H. and Khoualdia, T. (2020) ‘Application of optimised neural networks models in gears and bearings faults diagnosis’, Int. J. Vehicle Noise and Vibration, Vol. 16, Nos. 1/2, pp.30–45.
1 Introduction

Faults occurring in gearbox components such as bearing and gears are a major problem in rotating machines and they are diagnosed mainly by vibration analysis. The vibrations produced by their faults are non-stationary in nature, so it is essential to obtain the fault features accurately (Muruganatham et al., 2013; Xiang and Gao, 2017). Bearing and gear defects must be detected as early as possible to avoid fatal breakdowns of machines and prevent loss of production and human casualties (Xiao et al., 2017).

Bearings are among the most common causes of failure in rotating machines, some works took bearing faults in their study, like Zhou et al. (2016); they introduced a method called shift-invariant dictionary learning as a feature extraction technique. The adaptive technique is applied to study both simulated and experiment signals with specific notch size; an approach based on this method and hidden Markov model is addressed for bearing fault diagnosis. Dolenc et al. (2016) proposed a diagnostic procedure for the detection of distributed bearing faults employing a vibration analysis method for diagnosis, an envelope spectra comparison of vibration signals is used to distinguish the localised faults of the distributed one.

As well, Golafshan and Sanliturk (2016) applied to the ball bearing vibration signals and also for their spectra, to eliminate the background noise, a method called Singular Value Decomposition and a matrix of Hankel, to improve the reliability of fault detection process. In the same scope, He et al. (2016) used in a gearbox a signal processing method of time and frequency-domain vibration signals from four states of cylindrical roller bearings and they conducted a comprehensive evaluation of the vibration characteristics.

Gears are important parts like bearings in gearboxes and diagnosis of their faults is very necessary. Using only the structural vibrational response of the system, Taktak et al. (2012) used a method called blind source separation (BSS) to determine the waveform and excitation frequency of internal faults present in rotating mechanisms. However to diagnose gears, Waqar and Demetgil (2016) propose a diagnosis technique of gear faults under different working conditions (healthy, faulty and oil levels), a classification method based on the multi layer perceptron of an artificial neural network (ANN), which gives the possibility to predict the speed and oil level of the gearbox and a possible fault diagnosis. Based on reliability and a developed analytical mathematical method, Khalil et
K. Khoualdia et al. (2012) predicted the remaining lifetime of cracked and defect of gear tooth in a gearbox under different operating conditions. El Morsy and Achtenová (2015) applied an autocorrelation analysis method in order to detect an artificial pitting faults in gearbox of vehicle. For gear crack diagnosis in wind turbines, Li et al. (2016) apply the bounded component analysis for vibration study to eliminate the noise and disturbance signal components, and then the autoregressive model with prior data about faults of the gear is used to monitor the fault vibration cause.

Other works studied planetary gearbox faults and applied many methods to diagnose its faults and the results are validated by experiment, like Feng et al. (2016), they derived an explicit equation for calculating the characteristic frequency fault of each part of the planetary bearing (the outer race, inner race and rolling element), and they also developed its vibration signal model for different fault cases. Liu et al. (2016) modelled all the vibration signals sources and transmission way effects for planetary gear set, the vibration way effects are modelled in the gearbox to the different cases to the transducer situations, then they analyse the influences of diverse transmission paths on resultant vibration signals.

The ANNs approach for faults analysis has already attracted much interest of several researchers. Among them, Khoualdia et al. (2016) developed an approach based on a multi-objective optimisation with the L27 Taguchi standard orthogonal array and grey relational analysis, to find the best architecture of neural networks for monitoring and diagnosis of rotating machinery. However, in the present study, an attempt has been made to develop a reliable methodology to improve classification performance in order to have a monitoring and diagnosis tool to detect gear and bearing faults with the best NNM. For this goal, a test rig for vibration data collection is installed and a design of experiment based on Taguchi method and a comparison between two learning algorithms (back-propagation of the gradient error associated with Levenberg-Marquardt and back-propagation of the gradient error with momentum) are used.

2 Experimental set up

The experiments were performed on an experimental test rig (Figure 1) built specifically for the diagnosis of some faults in rotating machinery: healthy, unbalance, bearings faults, gears faults and that the combination of these states. In our study we focused on gears and bearings failures.

The test rig is fixed on the table of a milling machine, which gives us a good stability, a good fit of the driving shaft with the test rig shaft. The speed range is between 45 rpm and 2000 rpm. The driving motor is set to rotate with three different speeds, the speeds of 355 rpm, 710 rpm and 1400 rpm, are taken for all tests. The test rig consists of helical gears, as shown in Figure 1, the gear (G2) can be replaced by either healthy or faulty state to simulate the operating conditions of different gear faults. The gear (G1) that is mounted on the driving shaft is meshing with the gear (G2). The milling machines drive with a rigid coupling the driving shaft. By four sets of ball bearings the shafts are supported, close to the worn bearing, the accelerometer (vib 6.142 R) was mounted on the casing. The data acquisition device Vibexpert (vib 5.300) recorded the vibration signals of different functioning cases.
Figure 1 The test rig with the data acquisition device Vibexpert (vib 5.300): (a) kinematic chains of the test rig and (b) picture of the test rig (see online version for colours)

Four states of bearing and gear that are illustrated in this experiment, first state represents the bearing and gear transmission system functioning under the healthy condition, faulty bearing and gear conditions are compared with this baseline of operation. In second state, the nearest bearing (B2) of the gear (G2) as shown in Figure 1(a), which will be worn later, has been worn by a hard powder introduced into the bearing between its different parts in order to create an abrasive wear, Figure 2(a) and (b). In the third state, tooth of the gear (G2) is entirely removed to simulate the gear with a broken tooth Figure 2(c). Parameters of spur gear teeth (G2) and ball bearing UC204 (B2) are illustrated respectively in Tables 1 and 2.

Figure 2 Pictures of bearings and gear faults: (a) outer race; (b) inner race and (c) broken tooth (see online version for colours)

Table 1 Experimental test rig spur gear teeth parameters

| Parameters                  | Values  |
|-----------------------------|---------|
| Number of teeth             | 29      |
| Outer diameter (mm)         | 70.14   |
| Root diameter (mm)          | 58.74   |
| Pitch circle (mm)           | 64.54   |
| Working depth (mm)          | 5.74    |
| Pressure angle              | 20°     |
| Centre distance (mm)        | 85.6    |
| Circular tooth thickness (mm)| 4.32  |
| Axial circular pitch (mm)   | 7.48    |
| Normal circular pitch (mm)  | 6.18    |
Table 2 Ball bearing UC204 parameters

| Parameters                          | Values |
|-------------------------------------|--------|
| Number of rolling elements          | 8      |
| Pitch circle diameter (mm)          | 29.14  |
| Bearing bore diameter (mm)          | 20     |
| Bearing outside diameter (mm)       | 47     |
| Roller diameter (mm)                | 7.88   |
| Bearing width (mm)                  | 6.76   |
| Inner race width (mm)               | 31     |
| Outer race width (mm)               | 16     |
| Characteristic frequencies in Hz at 60 RPM |         |
| Cage frequency                      | 0.382  |
| Rolling element frequency           | 4.036  |
| Outer race frequency                | 3.063  |

3 Experimental analysis

3.1 Feature extraction

In order to illustrate the traditional FFT method example plots are shown in this section. The time and the frequency representation of the vibration high frequency signal with three different selected speeds, the speeds of 355 rpm, 710 rpm and 1400 rpm, in radial and axial directions simultaneously, are shown in Figure 3. For 1035 gear torque ratio, the speed of 355 rpm becomes 367.25 rpm, the same thing is for 710 rpm and 1400 rpm speeds (Kuang and Lin, 2003). The signals in time domain have elevated amplitude in the cases of faults compared with the normal state, the elevation of the amplitude is explained by the flaking of the bearing, a modulation of the signal in the case of the gears and for the combined case we can see more peaks of amplitude.

Figure 3 Acquired signals and their frequency spectrum diagrams for different faults states: (a) acquired signals and (b) FFT of acquired signals (see online version for colours)
However, in frequency domain signal for the different experimental bearings and gears and both of these states, in three different speeds and two directions (axial and radial), there is a clear difference, so the figures show a variation of amplitude level according to the increasing speeds.

There is a clearly higher peak of frequencies in each case corresponding to the kind of fault. The fault frequencies given in Table 3 are calculated to determine the corresponding peaks in frequency domain. Examples of spectra in frequency domain are illustrated in Figure 4.

**Figure 4** Frequency spectra of normal and faulty states for different speeds and directions: (a) speed of 335 rpm and b direction; (b) speed of 710 rpm and b direction; (c) speed of 1400 rpm and a direction and (d) speed of 1400 rpm and b direction (see online version for colours)

3.2 Time domain features

For each event of operation we recorded vibration signals with multiple repetitions for each measure with the speeds of 355 rpm, 710 rpm and 1400 rpm, in two different directions. To have enough individuals in the various operating modes (with and without fault) and after a preliminary analysis, we have chosen to calculate the following principal time indicators: RMS value, mean square value, variance, kurtosis factor, crest factor and clearance factor or margin factor, as given in Table 4 (Ali et al., 2015; Laissaoui et al., 2018; Dhamande and Chaudhari, 2018).
Table 3  Different frequencies corresponding to each kind of fault

| Different components | Faults frequency of the three selected speed |
|----------------------|---------------------------------------------|
|                      | $v_1 = 367.25$ rpm | $v_2 = 734.49$ rpm | $v_3 = 1448.3$ rpm |
| The shaft frequency (fr) in Hz | 6.12 | 12.24 | 24.13 |
| Outer race frequency (fo) in Hz | 18.75 | 37.50 | 73.92 |
| Inner race frequency (fi) in Hz | 30.21 | 60.42 | 119.10 |
| Cage frequency (fc) in Hz | 2.34 | 4.69 | 9.24 |
| Ball of the bearing frequency (fb) in Hz | 23.14 | 46.28 | 91.23 |
| Gear frequency (fg) in Hz | 177.48 | 354.96 | 699.77 |

Table 4  Time domain features

| Feature | Equation | Definition |
|---------|----------|------------|
| Maximum value | $\max \left| x_i \right|$ | Max value of the signal in the time domain |
| RMS value | $\sqrt{\frac{1}{N} \sum_{i=1}^{N} x_i^2}$ | Root mean square of the signal |
| Mean square value | $\text{MSV} = \frac{1}{N} \sum_{i=1}^{N} x_i^2$ | Mean square of the signal |
| Variance | $\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \text{MSV})^2$ | Variance of the signal |
| Kurtosis factor | $\frac{\left( N \sum_{i=1}^{N} (x_i - \overline{x})^4 \right)}{\left( \sum_{i=1}^{N} (x_i - \overline{x})^2 \right)^2}$ | Fourth normalised moment of the signal |
| Crest factor | $\frac{\max |x_i|}{\text{RMS}}$ | Ratio of maximum amplitude to RMS |
| Clearance factor | $\frac{\left( \max |x_i| \right)}{\sqrt{\frac{1}{N} \sum_{i=1}^{N} |x_i|^2}}$ | Ratio of maximum amplitude to the mean value |

4  The artificial neural network models

4.1 The optimisation of the MLP

Inspired by and based on the structure and operation of a biological neural network, an ANN is modelled. A multi-layer perceptron (MLP) network training with back-propagation algorithm is one of the commonest applications of ANN (Yuce et al., 2016; Sanz et al., 2012).

Our network MLP consists of three layers, as shown in Figure 5, input layer, output layer (which are controlled, because they have a fixed number of neurons) and a hidden layer that is uncontrolled, because we must vary and train it to obtain the best neural model. The input layer is composed of the main indicators used in vibratory analysis; the indicators are shown in Table 4. Though in the hidden layer, we must find the number
(NHL), the number of neurons in each hidden layer (NNHL) and the activation functions (tansig, logsig and purelin), for the case of the second learning algorithm, back-propagation of the gradient error with momentum, we have two more parameters (learning rate and momentum) to be varied, and finally our output layers are composed of a fixed number of neurons represented by a codification of the different conditions of operation, healthy, bearing faults, gears faults and combined defects of gear and bearing, for three different speeds and two directions (axial and radial).

**Figure 5**  The proposed topology of ANN

In the case of the back-propagation of the gradient error associated with Levenberg-Marquardt algorithm the number of the possible combinations for three levels and five factors is \(3^5 = 243\), this number that should increase with the learning repetition of each combination, thing that is very difficult to achieve without errors.

For the second learning algorithm, back-propagation of the gradient error with momentum, the number of combination becomes even more important, because with three levels and seven factors the number of combinations is equal to 2187, number that increases with the learning repetition of each case. The different combinations of our neural network are showed in Table 5.

**Table 5**  Independent variables and their levels of the two training algorithm (Levenberg-Marquardt algorithm and momentum algorithm)

| Levels | Number of hidden layers (NHL) | Number of neurons in the hidden layer (NNHL) | Learning rate (Lr) | Momentum (Mc) | Activation functions |
|--------|-----------------------------|---------------------------------------------|------------------|---------------|---------------------|
| 1      | 1                           | 4                                           | 0.1              | 0.1           | tansig              |
| 2      | 2                           | 13                                          | 0.5              | 0.5           | logsig              |
| 3      | 4                           | 22                                          | 0.9              | 0.9           | purelin             |

Lr, Mc: Addition parameters of Momentum algorithm.

However, for a minimum of combining networks and with the objective to have the best combination, an optimisation method of the hidden layers is used in order to have the best
learning network. We have applied the Taguchi method which serves to reduce the number of combinations for each learning algorithm (Sahu and Pal, 2015) (back-propagation of the gradient error associated with Levenberg-Marquardt and back-propagation of the gradient error with momentum) (Saravanan and Ramachandran, 2010).

4.2 Taguchi methods

Taguchi method can economically solve industry’s problems (Chen et al., 2016). According to our algorithms (back-propagation of the gradient error associated with Levenberg-Marquardt and back-propagation of the gradient error with momentum), for three levels and five factors in the first one, three levels seven factors in the second and referring to Taguchi tables, the most appropriate format is the L27 orthogonal array (Pandiarajan et al., 2016). The selected format is used to maximise the correlation parameters, to minimise the mean squared error (MSE) and number of epochs in order to find the best neural network model (NNM) architecture. Taguchi designs recognise that not all factors that cause variability can be controlled. These uncontrollable factors are called noise factors and try to identify controllable factors that minimise the effect of the noise factors signal to noise ratio (SNR). During the learning of the neural networks, we manipulate control factors to evaluate variability that occurs and then determine optimal control factor settings, which minimise the number of variability of combination in hidden layer and are also used to measure the significant factors through the ANOVA (Huang and Yu, 2016). Taguchi quality characteristic is such as smaller the better, higher the better type or nominal the best (Xie and Yuan, 2016; Nor et al., 2011), the cases with respective equations are presented as following:

Higher is the better calculated as follows:

\[
\frac{S}{N} \text{ ratio } = \eta_j = -10 \log \left( \frac{1}{n} \sum_{i=1}^{n} \frac{1}{y_{ij}} \right)
\]  

where \( y_{ij} \) = ith replicate of jth response, \( n \) = number of repetitions = 1, 2, ..., \( n \); \( j = 1, 2, ..., k \). Equation (1) is used to where maximisation of the quality characteristic of interest is desired.

Smaller is the better:

\[
\frac{S}{N} \text{ ratio } = \eta_j = -10 \log \left( \frac{1}{n} \sum_{i=1}^{n} y_{ij} \right)
\]  

Equation (2) is applied where minimisation of the quality characteristic is desired.

Nominal is the best:

\[
\frac{S}{N} \text{ ratio } = \eta_j = 10 \log \left( \frac{\Sigma_{j=1}^{n} \bar{y}_j}{s_j^2} \right) = 10 \log \left( \frac{\Sigma_{j=1}^{n} \bar{y}_j}{s_j^2} \right)
\]  

where \( \bar{y} = \frac{y_1 + y_2 + ... + y_n}{n} \) and \( \bar{y} = \frac{\Sigma_{j=1}^{n} \bar{y}_j}{s_j^2} \).

The nominal is the best type for minimisation of the MSE around a specific objective value (Nor et al., 2011).
However, the process is designed with the goal to improve the best model of our neural networks. Taguchi designs recognise that not all factors that cause variability can be controlled.

4.3 The analysis of variance

The analysis of variance (ANOVA) is performed to see which process parameters are statistically significant. With the SNR and ANOVA analyses, the optimal combination of the process parameters can be predicted.

SNR can be effectively employed to obtain the optimum level of each parameter and set of optimum parametric condition (to maximise) so to have the best of correlation parameters, the MSE and the number of epochs for this goal the overall SNR ratio for the model was obtained from the ANOVA study (Farbod et al., 2016; Dhawane et al., 2016; Varala et al., 2016).

4.4 Results and discussion

As shown in Tables 6 and 7, an orthogonal design is selected to perform the training combinations and to calculate the factor effects, Taguchi method is used to carry out the analysis and provide a signal-to-noise ratio for each of the factors of our two learning algorithms. The signal-to-noise ratio, according to the criterion ‘Higher is better’, is expressed with equation (1) (Yuce et al., 2014).

Table 6 
Taguchi orthogonal table: simulation using a neural network with a training algorithm type (back-propagation of the gradient error associated with Levenberg-Marquardt algorithm

| Learning No. | The 27 Taguchi combinations | Correlation parameters, mean squared error and no. of epochs |
|--------------|------------------------------|-------------------------------------------------------------|
|              | NHL | NNHL | AFIL | AFHL | AFOL | RA  | RV  | RG  | MSE | Epochs |
| 1            | 1   | 1    | 1    | 1    | 1    | 0.9747 | 0.9574 | 0.9695 | 0.0360 | 29 |
| 2            | 1   | 1    | 1    | 1    | 2    | 0.1062 | 0.1023 | 0.1040 | 1.2618 | 25 |
| 3            | 1   | 1    | 1    | 1    | 3    | 0.7259 | 0.8558 | 0.7681 | 0.3681 | 30 |
| 4            | 1   | 2    | 2    | 2    | 1    | 0.6001 | 0.5585 | 0.5879 | 0.6291 | 41 |
| 5            | 1   | 2    | 2    | 2    | 2    | 0.0559 | 0.0889 | 0.0644 | 1.1726 | 9  |
| 6            | 1   | 2    | 2    | 2    | 3    | 0.9376 | 0.9419 | 0.9373 | 0.1023 | 33 |
| 7            | 1   | 3    | 3    | 3    | 1    | 0.8896 | 0.8134 | 0.8666 | 0.1528 | 31 |
| 8            | 1   | 3    | 3    | 3    | 2    | 0.5726 | 0.3934 | 0.5172 | 0.8951 | 16 |
| 9            | 1   | 3    | 3    | 3    | 3    | 0.7248 | 0.3004 | 0.5543 | 0.3746 | 11 |
| 10           | 2   | 1    | 2    | 3    | 1    | 0.2053 | 0.1152 | 0.1789 | 1.3066 | 19 |
| 11           | 2   | 1    | 2    | 3    | 2    | 0.1464 | 0.1676 | 0.1497 | 1.6327 | 12 |
| 12           | 2   | 1    | 2    | 3    | 3    | 0.4876 | 0.4144 | 0.4612 | 0.5544 | 14 |
| 13           | 2   | 2    | 3    | 1    | 1    | 0.2773 | 0.1405 | 0.2357 | 1.1550 | 29 |
| 14           | 2   | 2    | 3    | 1    | 2    | 0.2195 | 0.3331 | 0.2551 | 1.1600 | 47 |
| 15           | 2   | 2    | 3    | 1    | 3    | 0.8244 | 0.9217 | 0.8523 | 0.2603 | 26 |
Table 6  Taguchi orthogonal table: simulation using a neural network with a training algorithm type (back-propagation of the gradient error associated with Levenberg-Marquardt algorithm (continued))

| Learning No. | The 27 Taguchi combinations | Correlation parameters, mean squared error and no. of epochs |
|--------------|------------------------------|-------------------------------------------------------------|
|              | NHL  | NNHL  | AFIL | AFHL | AFOL | RA    | RV    | RG    | MSE   | Epochs |
| 16           | 2    | 3     | 1    | 2    | 1    | 0.8945 | 0.9049 | 0.8989 | 0.1428 | 54     |
| 17           | 2    | 3     | 1    | 2    | 2    | 0.1289 | 0.1682 | 0.1457 | 0.9937 | 9      |
| 18           | 2    | 3     | 1    | 2    | 3    | 0.9984 | 0.8285 | 0.9463 | 0.0018 | 59     |
| 19           | 3    | 1     | 3    | 2    | 1    | 0.1218 | 0.0192 | 0.0808 | 1.5136 | 55     |
| 20           | 3    | 1     | 3    | 2    | 2    | 0.0658 | 0.0294 | 0.0547 | 1.4316 | 8      |
| 21           | 3    | 1     | 3    | 2    | 3    | 0.5864 | 0.8109 | 0.6536 | 0.5556 | 11     |
| 22           | 3    | 2     | 1    | 3    | 1    | 0.6202 | 0.7566 | 0.6573 | 0.5389 | 24     |
| 23           | 3    | 2     | 1    | 3    | 2    | 0.1484 | 0.0032 | 0.1082 | 1.2682 | 90     |
| 24           | 3    | 2     | 1    | 3    | 3    | 0.7486 | 0.7212 | 0.7409 | 0.2942 | 19     |
| 25           | 3    | 3     | 2    | 1    | 1    | 0.9079 | 0.8285 | 0.8840 | 0.1118 | 31     |
| 26           | 3    | 3     | 2    | 1    | 2    | 0.0698 | 0.2277 | 0.1200 | 1.0691 | 29     |
| 27           | 3    | 3     | 2    | 1    | 3    | 0.9516 | 0.9200 | 0.9397 | 0.0746 | 24     |

Table 7  Taguchi orthogonal table: simulation using a neural network with a training algorithm type back-propagation of the gradient error associated with momentum algorithm

| Learning No. | The 27 Taguchi combinations | Correlation parameters, mean squared error and no. of epochs |
|--------------|------------------------------|-------------------------------------------------------------|
|              | NHL  | NNHL  | Lr   | Mc   | AFIL | AFHL | AFOL | RA    | RV    | RG    | MSE   | IT   |
| 1            | 1    | 1     | 1    | 1    | 1    | 1    | 1    | 0.0666 | 0.0523 | 0.0604 | 1.7979 | 1761 |
| 2            | 1    | 1     | 1    | 1    | 2    | 2    | 2    | 0.0963 | 0.0996 | 0.0759 | 1.3303 | 7089 |
| 3            | 1    | 1     | 1    | 1    | 3    | 3    | 3    | 0.1716 | 0.1731 | 0.1733 | 0.8273 | 1831 |
| 4            | 1    | 2     | 2    | 1    | 1    | 1    | 1    | 0.6284 | 0.5154 | 0.5950 | 0.5197 | 9850 |
| 5            | 1    | 2     | 2    | 2    | 2    | 2    | 2    | 0.1006 | 0.1481 | 0.1223 | 1.146  | 6996 |
| 6            | 1    | 2     | 2    | 3    | 3    | 3    | 3    | 0.5654 | 0.5444 | 0.5589 | 0.5765 | 7123 |
| 7            | 1    | 3     | 3    | 3    | 1    | 1    | 1    | 0.1771 | 0.0425 | 0.1149 | 1.2390 | 6205 |
| 8            | 1    | 3     | 3    | 3    | 2    | 2    | 2    | 0.3022 | 0.1261 | 0.0587 | 1.7035 | 6840 |
| 9            | 1    | 3     | 3    | 3    | 3    | 3    | 3    | 0.2119 | 0.2455 | 0.2214 | 0.8209 | 1093 |
| 10           | 1    | 2     | 3    | 1    | 2    | 3    | 1    | 0.1762 | 0.1457 | 0.1642 | 0.8331 | 2057 |
| 11           | 2    | 1     | 2    | 3    | 2    | 3    | 1    | 0.1878 | 0.1093 | 0.1636 | 1.1033 | 8298 |
| 12           | 2    | 1     | 2    | 3    | 3    | 1    | 1    | 0.1159 | 0.2425 | 0.1497 | 1.4703 | 7281 |
| 13           | 2    | 2     | 3    | 1    | 1    | 2    | 3    | 0.2108 | 0.2002 | 0.2067 | 0.8098 | 1310 |
| 14           | 2    | 2     | 3    | 1    | 2    | 3    | 1    | 0.1557 | 0.1111 | 0.1408 | 1.0121 | 1143 |
| 15           | 2    | 2     | 3    | 1    | 3    | 1    | 1    | 0.2626 | 0.2901 | 0.2719 | 1.0327 | 7491 |
| 16           | 2    | 3     | 1    | 2    | 1    | 2    | 3    | 0.2226 | 0.1611 | 0.2036 | 0.8049 | 7162 |
| 17           | 2    | 3     | 1    | 2    | 2    | 3    | 1    | 0.2935 | 0.2017 | 0.2678 | 0.9164 | 1708 |
Table 7  
| The 27 Taguchi combinations | Correlation parameters, mean squared error and no. of epochs |
|-----------------------------|------------------------------------------------------------|
| L  | NHL | NNHL | Lr | Mc | AFIL | AFHL | AFOL | RA  | RV  | RG  | MSE   | IT   |
|----|-----|------|----|----|------|------|------|-----|-----|-----|-------|------|
| 18 | 2   | 3    | 1  | 2  | 3    | 1    | 2    | 0.1244 | 0.1104 | 0.1226 | 1.3492 | 5172 |
| 19 | 3   | 1    | 3  | 2  | 2    | 1    | 3    | 0.1895 | 0.3188 | 0.2316 | 1.2130 | 6775 |
| 20 | 3   | 1    | 3  | 2  | 2    | 1    | 3    | 0.2267 | 0.2219 | 0.2239 | 0.8054 | 3234 |
| 21 | 3   | 1    | 3  | 2  | 3    | 2    | 1    | 0.1755 | 0.1785 | 0.1761 | 1.1211 | 6854 |
| 22 | 3   | 2    | 1  | 3  | 2    | 1    | 3    | 0.1896 | 0.2316 | 0.2076 | 1.1633 | 1217 |
| 23 | 3   | 2    | 1  | 3  | 3    | 2    | 1    | 0.1640 | 0.2464 | 0.1922 | 0.8437 | 1498 |
| 24 | 3   | 3    | 2  | 1  | 1    | 1    | 3    | 0.3099 | 0.3097 | 0.3096 | 0.8136 | 6528 |
| 25 | 3   | 3    | 2  | 1  | 1    | 2    | 1    | 0.0872 | 0.1127 | 0.0951 | 1.7499 | 2135 |
| 26 | 3   | 3    | 2  | 1  | 3    | 2    | 1    | 0.2205 | 0.2019 | 0.2136 | 0.8059 | 1364 |
| 27 | 3   | 3    | 2  | 3  | 1    | 3    | 2    | 0.1635 | 0.2945 | 0.2000 | 1.3005 | 4996 |

According to the analysis of variance ANOVA the results, given respectively in Figures 6 and 7, are: The best combination of our ANN model, in the case of the back-propagation of the gradient error associated with Levenberg-Marquardt, is NHL = 3, NNHL = 3, AFIL = 2, AFHL = 1 and AFOL = 3. However, for the second algorithm, back-propagation of the gradient error with momentum, the best combination of the ANNs model is NHL = 3, NNHL = 3, Lr = 2, Mc = 1, AFIL = 3, AFHL = 1 and AFOL = 3.

Figure 6  
Response graph of the algorithm back-propagation of the gradient error associated with Levenberg-Marquardt (see online version for colours)
After the training of our selected MLP, in agreement with the combinations given by the analysis of variance ANOVA, we have the following results:

- For ANNs with back-propagation of the gradient error associated with Levenberg-Marquardt: test correlation (RA) = 0.9988, validation of correlation (RV) = 0.9158, global correlation (RG) = 0.9734, mean square error (MSE) = 0.0001 and number of epochs (IT) for only 17 iterations.

- For ANNs with back-propagation of the gradient error with momentum: test correlation (RA = 0.9197, validation of correlation (RV = 0.8491, global correlation (RG = 0.8971, mean square error (MSE = 0.0832 and number of epochs (IT) = iterations.

The performances of the best MLP combinations are illustrated in Figures 8 and 9.
5 Conclusion

In order to bring reliable methodologies, two back-propagation algorithms of MLP network training, the back-propagation of the gradient error associated with Levenberg-Marquardt and back-propagation of the gradient error with momentum, are compared. Experimental and statistical methods are used to improve the classification performances. The target of this study is monitoring and diagnosis for early detection of the rotating machinery faults, in particular bearing, gears and combined (gears-bearing) faults. To determine the best MLP, the Taguchi approach is utilised to optimise the two algorithms before their comparison.

The optimal structure of the intelligent system in the case of the back-propagation of the gradient error associated with Levenberg-Marquardt, $NHL = 3, \text{NNHL} = 3, AFI = 2, AFHL = 1$ and $AFOL = 3$, reached quickly with only 17 epochs the desired convergence, the mean square error is equal to 0,0001 and the correlation coefficients approximated 100%.

The optimal structure of the second intelligent system in the case of the back-propagation of the gradient error with momentum, $NHL = 3, \text{NNHL} = 3, Lr = 2, Mc = 1, AFIL = 3, AFHL = 1$ and $AFOL = 3$, reached the desired convergence after 6055 iterations, the mean square error is equal to 0,0832.

The total training time, of the MLP with back-propagation of the gradient error associated with Levenberg-Marquardt, does not exceed a few seconds which is very short compared to the MLP with back-propagation of the gradient error with momentum.
The presented method may provide practical, effective utilities for bearing and gear faults diagnosis and both of them as well as other industrial combinations of faults of complex machines.

References

Ali, J.B., Fnaiech, N., Saidi, L., Chebel-Morello, B. and Fnaiech, F. (2015) ‘Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals’, *Applied Acoustics*, Vol. 89, pp.16–27.

Chen, P.C., Yang, M.W., Wei, C.H. and Lin, S.Z. (2016) ‘Selection of blended amine for CO2 capture in a packed bed scrubber using the Taguchi method’, *International Journal of Greenhouse Gas Control*, Vol. 45, pp.245–252.

Dhamande, L.S. and Chaudhari, M.B. (2018) ‘Compound gear-bearing fault feature extraction using statistical features based on time-frequency method’, *Measurement*, Vol. 125, pp.63–77.

Dhawane, S.H., Kumar, T. and Halder, G. (2016) ‘Biodiesel synthesis from Hevea brasiliensis oil employing carbon supported heterogeneous catalyst: optimization by Taguchi method’, *Renewable Energy*, Vol. 89, pp.506–514.

Dolenc, B., Boškoski, P. and Juričić, Đ. (2016) ‘Distributed bearing fault diagnosis based on vibration analysis’, *Mechanical Systems and Signal Processing*, Vol. 66, pp.521–532.

El Morsy, M. and Achtenová, G. (2015) ‘Value of autocorrelation analysis in vehicle gearbox fault diagnosis’, *International Journal of Vehicle Noise and Vibration*, Vol. 11, No. 2, pp.165–184.

Farbod, M., Rafati, Z. and Shoushtari, M.Z. (2016) ‘Optimization of parameters for the synthesis of Y2Cu2O5 nanoparticles by Taguchi method and comparison of their magnetic and optical properties with their bulk counterpart’, *Journal of Magnetism and Magnetic Materials*, Vol. 407, pp.266–271.

Feng, Z., Ma, H. and Zuo, M.J. (2016) ‘Vibration signal models for fault diagnosis of planet bearings’, *Journal of Sound and Vibration*, Vol. 370, pp.372–393.

Golafshan, R. and Sanlurk, K.Y. (2016) ‘SVD and Hankel matrix based de-noising approach for ball bearing fault detection and its assessment using artificial faults’, *Mechanical Systems and Signal Processing*, Vol. 70, pp.36–50.

He, K., Liu, X. and Li, X. (2016) ‘Experimental study of the vibration characteristics of the fault cylindrical roller bearing of special vehicle gearbox’, *International Journal of Vehicle Noise and Vibration*, Vol. 12, No. 3, pp.241–259.

Huang, C.N. and Yu, C.C. (2016) ‘Integration of Taguchi’s method and multiple-input, multiple-output ANFIS inverse model for the optimal design of a water-cooled condenser’, *Applied Thermal Engineering*, Vol. 98, pp.605–609.

Khalil, M.I., El-morsy, M.S.A. and Abouel-seoud, S.A. (2012) ‘Optimisation of gearbox replacement policy using vibration measurement data’, *International Journal of Vehicle Noise and Vibration*, Vol. 8, No. 4, pp.337–351.

Khoualdia, T., Hadjadj, A.E., Bouacha, K. and Abdeslam, D.O. (2016) ‘Multi-objective optimization of ANN fault diagnosis model for rotating machinery using grey rational analysis in Taguchi method’, *The International Journal of Advanced Manufacturing Technology*, Vol. 89, Nos. 9–12, pp.3009–3020.

Kuang, J.H. and Lin, A.D. (2003) ‘Theoretical aspects of torque responses in spur gearing due to mesh stiffness variation’, *Mechanical Systems and Signal Processing*, Vol. 17, No. 2, pp.255–271.

Laissaoui, A., Bouzouane, B., Miloudi, A. and Hamzaoui, N. (2018) ‘Perceptive analysis of bearing defects (Contribution to vibration monitoring)’, *Applied Acoustics*, Vol. 140, pp.248–255.
Li, Z., Yan, X., Wang, X. and Peng, Z. (2016) ‘Detection of gear cracks in a complex gearbox of wind turbines using supervised bounded component analysis of vibration signals collected from multi-channel sensors’, *Journal of Sound and Vibration*, Vol. 371, pp.406–433.

Liu, L., Liang, X. and Zuo, M.J. (2016) ‘Vibration signal modeling of a planetary gear set with transmission path effect analysis’, *Measurement*, Vol. 85, pp.20–31.

Muruganatham, B., Sanjith, M.A., Krishnakumar, B. and Murty, S.S. (2013) ‘Roller element bearing fault diagnosis using singular spectrum analysis’, *Mechanical Systems and Signal Processing*, Vol. 35, Nos. 1–2, pp.150–166.

Nor, N.M., Muhamad, N., Ibrahim, M.H.I., Ruzi, M. and Jamaludin, K.R. (2011) ‘Optimization of injection molding parameter of Ti-6Al-4V powder mix with palm stearin and polyethylene for the highest green strength by using Taguchi method’, *International Journal of Mechanical and Materials Engineering*, Vol. 6, No. 1, pp.126–132.

Pandiarajan, S., Kumaran, S.S., Kumaraswamidhas, L.A. and Saravanan, R. (2016) ‘Interfacial microstructure and optimization of friction welding by Taguchi and ANOVA method on SA 213 tube to SA 387 tube plate without backing block using an external tool’, *Journal of Alloys and Compounds*, Vol. 654, pp.534–545.

Sahu, P.K. and Pal, S. (2015) ‘Multi-response optimization of process parameters in friction stir welded AM20 magnesium alloy by Taguchi grey relational analysis’, *Journal of Magnesium and Alloys*, Vol. 3, No. 1, pp.36–46.

Sanz, J., Perera, R. and Huerta, C. (2012) ‘Gear dynamics monitoring using discrete wavelet transformation and multi-layer perceptron neural networks’, *Applied Soft Computing*, Vol. 12, No. 9, pp.2867–2878.

Saravanan, N. and Ramachandran, K.I. (2010) ‘Incipient gear box fault diagnosis using discrete wavelet transform (DWT) for feature extraction and classification using artificial neural network (ANN)’, *Expert Systems with Applications*, Vol. 37, No. 6, pp.4168–4181.

Taktak, M., Toumi, D., Akrout, A., Abbès, M.S. and Haddar, M. (2012) ‘One stage spur gear transmission crankcase diagnosis using the independent components method’, *International Journal of Vehicle Noise and Vibration*, Vol. 8, No. 4, pp.387–400.

Varala, S., Kumari, A., Dharamija, B., Bhragava, S.K., Parthasarathy, R. and Satyavathi, B. (2016) ‘Removal of thorium (IV) from aqueous solutions by deoiled karanja seed cake: optimization using Taguchi method, equilibrium, kinetic and thermodynamic studies’, *Journal of Environmental Chemical Engineering*, Vol. 4, No. 1, pp.405–417.

Waqar, T. and Demetgul, M. (2016) ‘Thermal analysis MLP neural network based fault diagnosis on worm gears’, *Measurement*, Vol. 86, pp.56–66.

Xiang, L. and Gao, N. (2017) ‘Coupled torsion–bending dynamic analysis of gear-rotor-bearing system with eccentricity fluctuation’, *Applied Mathematical Modelling*, Vol. 50, pp.569–584.

Xiao, H., Zhou, X., Liu, J. and Shao, Y. (2017) ‘Vibration transmission and energy dissipation through the gear-shaft-bearing-housing system subjected to impulse force on gear’, *Measurement*, Vol. 102, pp.64–79.

Xie, J. and Yuan, C. (2016) ‘Parametric study of ice thermal storage system with thin layer ring by Taguchi method’, *Applied Thermal Engineering*, Vol. 98, pp.246–255.

Yuce, B., Mastrocinque, E., Packianather, M.S., Pham, D., Lambiase, A. and Fruggiero, F. (2014) ‘Neural network design and feature selection using principal component analysis and Taguchi method for identifying wood veneer defects’, *Production and Manufacturing Research*, Vol. 2, No. 1, pp.291–308.

Yuce, B., Rezgui, Y. and Moursедh, M. (2016) ‘ANN–GA smart appliance scheduling for optimised energy management in the domestic sector’, *Energy and Buildings*, Vol. 111, pp.311–325.

Zhou, H., Chen, J., Dong, G. and Wang, R. (2016) ‘Detection and diagnosis of bearing faults using shift-invariant dictionary learning and hidden Markov model’, *Mechanical Systems and Signal Processing*, Vol. 72, pp.65–79.