25th ABCM International Congress of Mechanical Engineering
October 20-25, 2019, Uberlândia, MG, Brazil

COB-2019-1630
STRUCTURAL HEALTH MONITORING IN MECHANICAL SYSTEMS FROM CHANGES IN THEIR VIBRATION CHARACTERISTICS BASED ON ARTIFICIAL NEURAL NETWORKS AND PCA

Luísa Rosenstock Völtz
Eduardo Lenz Cardoso
Ricardo De Medeiros
Santa Catarina State University, Department of Mechanical Engineering, Rua Paulo Malschitzki, 200. Zona Industrial Norte, 89219-710 Joinville, SC, Brazil.
luisa.voltz@gmail.com, eduardo.cardoso@udesc.br, ricardo.medeiros@udesc.br

Abstract. *Structural Health Monitoring (SHM) methods comprise early detection of modifications in a given set of characteristics of a system to predict failure. Inspired by the biological nerve system, artificial neural networks have been applied in machine learning for data classification and pattern recognition and can handle a wide variety of non-linear complex systems. This work proposes the use of both principal component analysis and feedforward artificial neural networks to analyze vibrational data to predict failure. To this end, three different mechanical systems: aluminum beams, rolling bearings, and composite plates were studied in details. For all systems, different damage scenarios are considered. The results show that the methodology can detect damages in the different mechanical systems applied with excellent levels of accuracy.*

Keywords: *Damage detection, Structural health monitoring (SHM), Artificial neural networks (ANNs), Principal component analysis (PCA), Vibration-based techniques.*

1. INTRODUCTION

Due to manufacturing imperfections, operational conditions, climatic variations, and minor accidents, structures can lose their full capacity to withstand the operational load, thus reducing residual life. Under these conditions, a continuous structural health monitoring is paramount to avoid loss of functionality. The most used technique for structural monitoring is a visual inspection, requiring full access to the structure, as well as being time-consuming and prone to human errors (Yuan, 2016).

Mechanical systems with rotating parts produce vibrations due to mechanical disturbances from different sources. Vibration is one of the features of mechanical machinery that if uncontrolled, can origin small or serious performance, operational or safety problems. Machine vibration is an essential element for Structural Health Monitoring (SHM), so there is a large interest in vibration-based techniques (Goyal and Pabla, 2016).

According to Bandara *et al.* (2014), vibration-based damage identification methods are considered to be very useful as they provide an on-line monitoring and global inspection mode. In this way, there is a trend of using frequency response functions (FRFs) for damage detection. The combination of FRFs with Artificial Neural Networks (ANNs) can provide a very promising tool for damage identification methodology, as shown in some works (Zang and Imregun, 2001; Li *et al.*, 2011; Samali *et al.*, 2012; Bandara *et al.*, 2014).

To overcome problems like a large number of inputs, when using large resolutions frequency data (FRF), some sort of data compression is needed. One established method in the literature is the Principal Component Analysis (PCA), providing not only data compression but assisting in feature extraction and conditioning (Bishop *et al.*, 1995).

The study performed by Li *et al.* (2011) investigated the use of the PCA and ANN to identify locations and severities of damage in steel beam structures from residual frequency response functions (FRFs). The learning process was done using back-propagation and conjugate gradient descent algorithm. Although the procedure was sensitive to sensor locations, the results showed the effectiveness of the joint use of PCA and ANN approach using vibration-based methods to damaged detection. Bandara *et al.* (2014) applied the same methodology, with FRFs from a finite element model of two steel-storey framed structured, PCA-compressed technique and ANN to locate both damage locations and severities. They measured FRFs from 14 different locations on the structure and then 14 backpropagation-ANNs were developed for each location. Single and multiple damage cases were also studied, showing good results for both cases. It was found that ANN trained with vertical FRFs measured from crossbeams showed a smaller error when compared to measurements from the horizontal direction. For a single damage case, in the testing set, the maximum mean-square error (MSE) was 3.04% for damage location and 2.58% for damage severity identification. For multiple damage scenarios, in the testing set, the maximum MSE was 4.46% for damage location and 3.57% for damage severities identification.

There are some distinct SHM-approaches used to detect damages in the structure, and it is broadly recognized that the success of a certain methodology may depend on the structure being monitored. Consequently, it is essential that new methodologies are verified against different mechanical systems and damage cases (Gul and Catbas, 2009). This work study the combined use of PCA and ANN, to damage detection in three different scenarios: metallic beams, rolling bearings, and composite plates. To this end, both concepts are briefly presented, and the experimental setup of all experiments are discussed in details. Finally, it is shown a discussion about the potentialities and limitations of the proposed methodology as a tool for supporting SHM systems design.

2. THEORETICAL BACKGROUND

This section presents the theoretical aspects of Principal Component Analysis (PCA) and Artificial Neural Networks (ANN).

2.1 Principal Component Analysis (PCA)

The main function of the PCA is to determine the coordinate axes of the new coordinate system considering the highest variance of the raw data. Through the PCA application, the raw data can be reduced in *P* components, called Principal Components (PCs). The PCs are extracted from the correlation matrix of the original data, through an eigenvalue and eigenvector problem (Zang and Imregun, 2001). Given a rectangular $M \times N$ matrix containing $M \times N$-dimensional real-valued vectors of data, $\mathbf{H}$, with a single element h_{ij} with *i* lines and *j* columns. The calculation of the mean response vector $\bar{h}$ is defined as

$$\bar{h_j} = \frac{\sum_{i=0}^{M} h_{ij}}{M}, \tag{1}$$

The standard deviation vector S^2 is defined as

$$S_j^2 = \frac{\sum_{i=0}^{M} (h_{ij} - \bar{h_j})^2}{M}. \tag{2}$$

A single element of the $\mathbf{H}$ matrix can than be replaced by

$$\tilde{h_{ij}}(\omega) = \frac{h_{ij} - \bar{h}_j}{S_j \sqrt{M}}, \tag{3}$$

The correlation matrix $C_{N \times N}$ is defined as

$$\mathbf{C}_{N \times N} = (\tilde{\mathbf{H}}_{M \times N})^T \tilde{\mathbf{H}}_{M \times N}. \tag{4}$$

The eigenvalues λ_k, k=1...N, of matrix $\mathbf{C}$ are known as PCs, such that

$$\mathbf{C}\phi_k = \lambda_k \phi_k, \tag{5}$$

where ϕ_k are the corresponding eigenvectors. By using a subset containing just the *P*-largest (most relevant) eigenvalues (P≤N), it is possible to reconstruct $\tilde{\mathbf{H}}$ with a smaller number of entries (columns), such that

$$\mathbf{A}_{M \times P} = \tilde{\mathbf{H}}_{M \times N} \phi_{N \times P}, \tag{6}$$

where $\mathbf{A}_{M \times P}$ retains the most significant features of the original matrix, and it is used as input to the neural network.

2.2 Artificial Neural Networks (ANNs)

According to Bishop *et al.* (1995), the goal of ANNs is to provide general non-linear parameterized mappings between a set of inputs and outputs variables. The most common form of ANN is the feedforward neural network, where this mapping is provided by layers of interconnected neurons. The importance of each connection is parametrized by weight and the sensitivity of each neuron depends on both threshold parameter (bias) and a nonlinear activation function. The most known activation function is defined as

$$\phi = \frac{1}{1 + e^{-v}}, \tag{7}$$

where *v* is the linear combiner between the weights and the inputs parameters, the logistic function assumes a continuous range of values from 0 to 1. The set of weights and bias, for the entire network, are the design variables of an optimization problem, aiming to minimize some Error or Cost function. One very common cost function is defined as

$$E(\mathbf{w},\mathbf{b}) = (\sum_{i=1}^{n}(z_i - y_i)^2)^{1/2} = \|\mathbf{z} - \mathbf{y}\|_2. \tag{8}$$

where **w** contains the weights, **b** the bias, *n* is the length of the output vector, **z** the target output and **y** the network output. The Eq. 8 is known as Euclidean distance (or L_2). Another type of the cost function is known as the Cross-Entropy function defined as (Bishop, 2006)

$$E(\mathbf{w},\mathbf{b}) = -\sum_{i=1}^{n}(z_i \ln y_i + (1 - z_i)ln(1 - y_i)). \tag{9}$$

According to Simard *et al.* (2003), the cross-entropy function trained faster than the quadratic function for classification problems.

The learning method is based on the gradient descent, and the weights and biases are iteratively adjusted, for *t* iterations, as follow (Haykin, 2007)

$$\Delta\mathbf{w}^t = \gamma\Delta\mathbf{w}^{t-1} + \alpha^t\mathbf{d}^t, \tag{10}$$

$$\Delta\mathbf{b}^t = \gamma\Delta\mathbf{b}^{t-1} + \alpha^t\mathbf{d}^t, \tag{11}$$

and

$$\mathbf{w}^t = \mathbf{w}^{t-1} + \Delta\mathbf{w}^t, \tag{12}$$

$$\mathbf{b}^t = \mathbf{b}^{t-1} + \Delta\mathbf{b}^t. \tag{13}$$

where **d** is the descent direction, γ is the momentum term and α is the step size. The steepest descent method is used in this work, and the sensitivity is found by using forward automatic differentiation (Revels *et al.*, 2016). The learning rate can be obtained using a fixed value or by using a line search technique during the learning process, such as Backtracking line search.

To overcome the overfitting problems, some regularization techniques can be adopted such as the L_1 regularization method applied to the quadratic cost function, defined as

$$E(\mathbf{w},\mathbf{b}) = \frac{1}{2n}\sum_{i=1}^{n}\|z_i - y_i\|^2 + \frac{\lambda}{2n}\sum_{w}|w| \tag{14}$$

where *E* is the error function defined previously, λ a control parameter that extent the penalty term and influence the solution and **w** are the all weights in the neural network (Nielsen, 2018).

According to Bishop *et al.* (1995), one way to control the overfitting involves the addition of artificial noise to the input data during training. Adding small amounts of artificial noise can increase generalization as long as the amount of noise is maintained small to have a little effect on the desired output. One technique consists of the addition of noise with Gaussian distribution to the uncorrelated input patterns (Da Silva and Adeodato, 2011). The expression used to generate the noise values is defined as

$$c = k * G(0,1) \tag{15}$$

where *c* is the noise added to every input variable, *k* is the quantity of noise, and *G* is the Gaussian distribution with zero mean and unitary variance.

3. DAMAGE DETECTION PROCEDURE

The methodology presented is validated in three cases using different mechanical systems. First, a vibration-based method is used, and the dynamic curves are taken. To reduce the data size, a PCA is applied, and the new data set is introduced in ANNs to damage detection.

3.1 Metallic Beams

Aluminum beams with a rectangular cross-section of 19.7mm *vs.* 4.9mm, and a length of 496mm, are studied with the damaged patterns: 2mm crack size, 4mm crack size, and 8mm crack size. The damages are performed in the middle of the beam. Vibration-based analyses are performed using Pulse LabShop software by Brüel & Kjaer, in free-free condition with an impact hammer (sensitivity 1.12 mV/N) and an uniaxial accelerometer (sensitivity 10 mV/g), as can be seen in Fig. 1. The excitation is applied in two different points (positions H_{11} at 8mm and H_{21} at 420mm), such that for each specimen two FRFs are measured. In the total 172 FRFs are taken from healthy (H) and damaged beams with 3201 frequency points from a frequency range of 0-1600 Hz (Völtz *et al.*, 2017). Figure 2 shows four FRFs from each case from one specimen. PCA is applied, and the data are reduced to 10 PCs, retained 98.04% of the total variance.

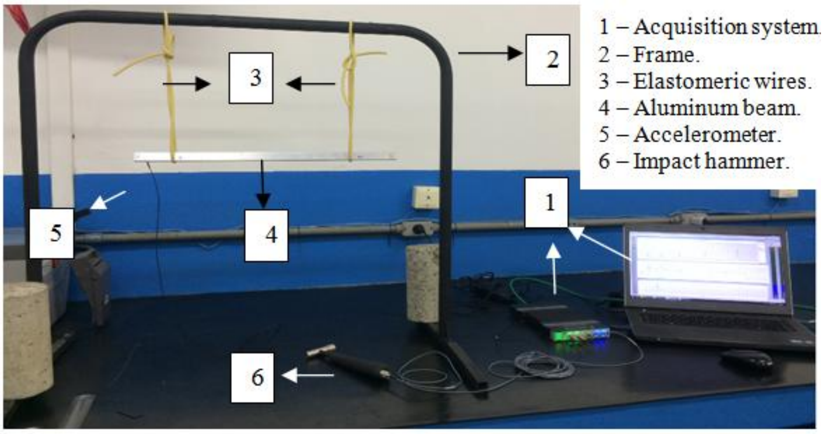

Figure 1. Experimental setup for aluminum beams (Völtz *et al.*, 2017).

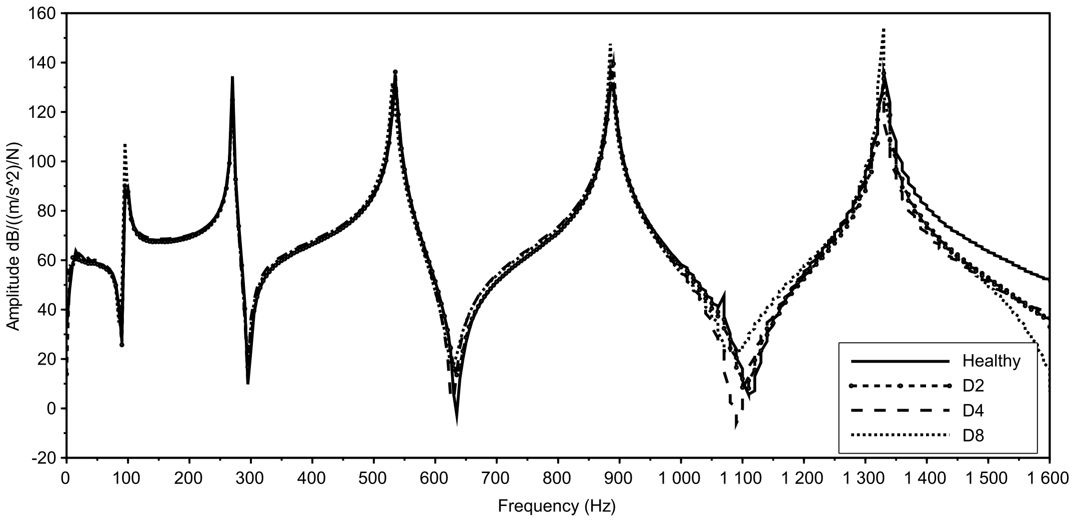

Figure 2. FRFs for healthy, damaged with 2mm crack size (D_2), damaged with 4mm crack size (D_4) and damaged with 8mm crack size (D_8) for H_{11} position.

The data is split into three sets (71% for the training, 10% for the validation and 19% for the testing sets). A fully connected multilayer perceptron neural network using the logistic function as activation function and quadratic function as a cost function is developed. The two patterns (healthy and damaged) are presented to the ANN for the learning process. The topology is ten inputs parameters, eight neurons in the first hidden layer, four neurons in the second hidden layer and two neurons in the output layer ([1,0] for healthy and [0,1] for damaged).

The momentum term is 0.8, and a fixed learning rate of 0.3 is applied. Also, 3% of the artificial noise is added to the training set and L_1 regularization technique is applied with a lambda value of 0.001. The ANN generalizes 93.55% of the pattern correctly, the results are summarized in Tab. 1. As can be observed there is presence of underfitting, since the accuracy of the training set is quite high. The low accuracy of the validation set can be explained by samples from this set is not being contained in the training set domain, or by the presence of atypical samples of specific pattern class.

3.2 Rolling Bearings

Two rolling bearings model NSK HR 32004 XJ are studied, by Zago (2016) and Zago (2018)). An apparatus developed by Zago (2016) is shown in Fig. 3, composed of a motor connected to an axis by a belt. An NSK 6006 DDU ball bearing

Table 1. ANNs results: aluminum beams.

Pattern cases	Training iterations	Training accuracy (%)	Validation accuracy (%)	Testing accuracy (%)
H+D	305	95.16	82.35	93.55

is used in the bushing closest to the motor, and an NSK HR 32004 XJ taper bearing is used in the bushing on the other end, which is the object of study. The applied load is composed of a mass of 1.3 kg in the center of the shaft, in order to apply a static load on the bearings. The undamaged and damaged Fast Fourier Transforms (FFTs) are obtained using an accelerometer (model 4397 by Brüel & Kjaer) attached to the bushing. The sampling frequency of 400 Hz was chosen to include more modes and to have the influence of the fault in the analysis range. Each measurement produces a time-amplitude data packet over 8 seconds. The damage is performed on the outer race surface of the rolling bearing (Zago, 2016). The undamaged and damage FFTs can be seen in Fig. 4.

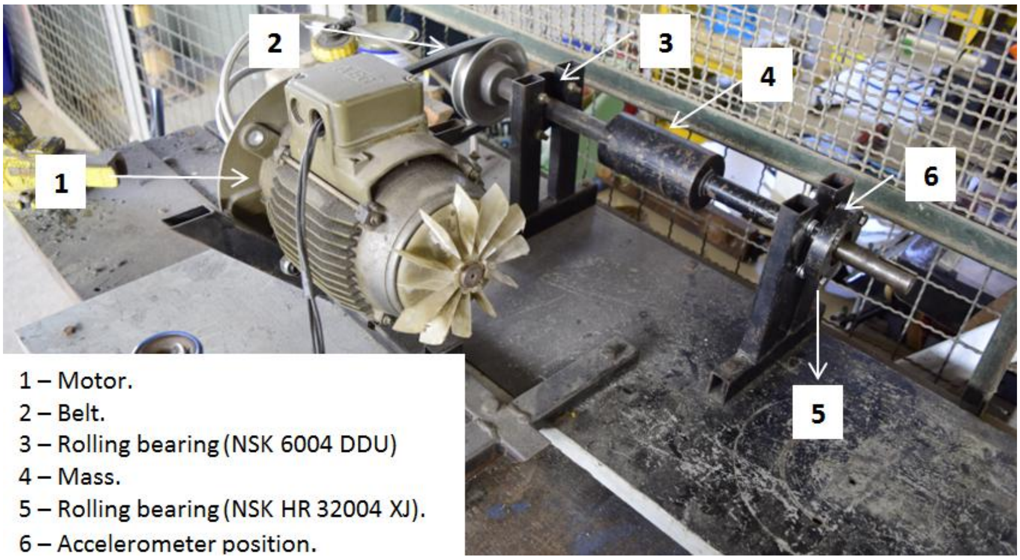

Figure 3. Experimental setup for rolling bearings. Adapted from Zago (2016).

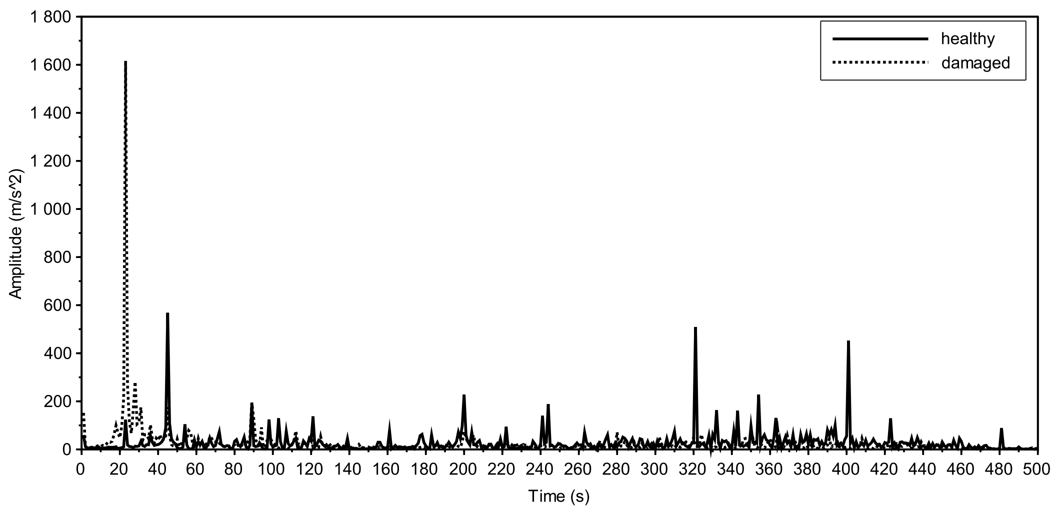

Figure 4. FFTs for healthy and damaged rolling bearing

PCA is applied to reduce the data. The data is also split into three sets (70% for the training set, 15% for the validation set, and 15% for the testing set). The author uses different values of the PCs and consequently the total variance to evaluate the behavior of the ANN. The logistic function and L_2 cost function is used. A momentum term of 0.1 and a fixed learning rate of 1.0 are used. The topology of the ANN is two hidden layers with twelve and ten neurons for the first and second hidden layers, respectively, and two output neurons.

From 37% to 80% of the total variance of the PCA, the ANN generalizes 100% of the patterns correctly, between undamaged and damaged patterns. The results are summarized in Tab. 2.

Table 2. Principal components, total variance and ANN results for each variance cases.

PC	Total variance (%)	Training iterations	Training error	Testing accuracy(%)
41	80	55	9.29E-9	100
23	70	53	4.26E-9	100
12	60	49	7.83E-9	100
6	51	56	4.13E-9	100
3	44	54	4.53E-9	100
2	37	55	4.86E-9	100

3.3 Composite Plates

Composite plates made of carbon fibers/epoxy resin are studied, with two stacking orientation $[0/15/\text{-}15/0/15/\text{-}15]_S$ and $[0]_8$. For the $[0]_8$ orientation group, five plates are damaged by impact loading, and one is damaged by drilling a center hole. For the $[0/15/\text{-}15/0/15/\text{-}15]_S$ plates, four are damaged by impact loading and two by delamination. The composites plates and the damaged procedures were executed previously by De Medeiros (2016), where the impact loading and the hole were performed in the center of the plates, on the other hand the delamination was done in specific spots. The FRFs are taken from healthy and damaged plates, in free-free condition using impact hammer, accelerometer (in 4 positions) and LMS acquisition system, as seen in Fig. 5. The analyzed frequency range is 0-1024 Hz with 2048 spectral points. Due to the interest of analyzing up to the fifth mode, and due the applicability of the Nyquist theorem it is necessary to analyze the frequency up to 1024 Hz. A total of 48 FRFs for $[0/15/\text{-}15/0/15/\text{-}15]_S$ and 56 FRFs for $[0]_8$ are taken (De Medeiros, 2016). The real part values of the FRFs and the accelerance values of the FRFs - $[0/15/\text{-}15/0/15/\text{-}15]_S$ stacking orientation - are shown in Fig. 7 and Fig. 6, respectively. Figure 8 represents the FRFs using accelerance values and Fig. 9 using real part values for $[0]_8$ stacking orientation.

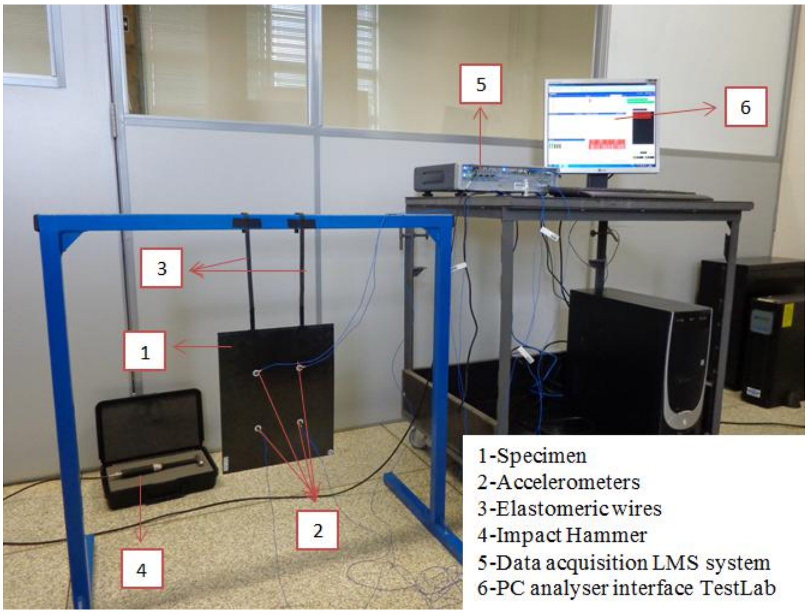

Figure 5. Experimental setup (De Medeiros, 2016).

The data is split into 62% for the training set, 21% for the validation set and 17% for the testing set. Logistic function, L_2 cost function, and cross-entropy cost function are used in the ANNs learning process. The topology and the parameters of each neural network simulation can be seen in Tab. 3. The PCA is applied to retain 99.19% and 98.55% using the first 20 PCs for accelerance values of the FRF and real values of the FRFs, respectively, from $[0]_8$ stacking group. The ANN generalizes 90% and 100% of the patterns correctly using accelerance values and real values, respectively. For $[0/15/\text{-}15/0/15/\text{-}15]_S$ stacking group, the PCA retained 98.11% of the total variance with 20 PCs using accelerance values of the FRFs and 98.85% using the first 30 PCs with the real part values.

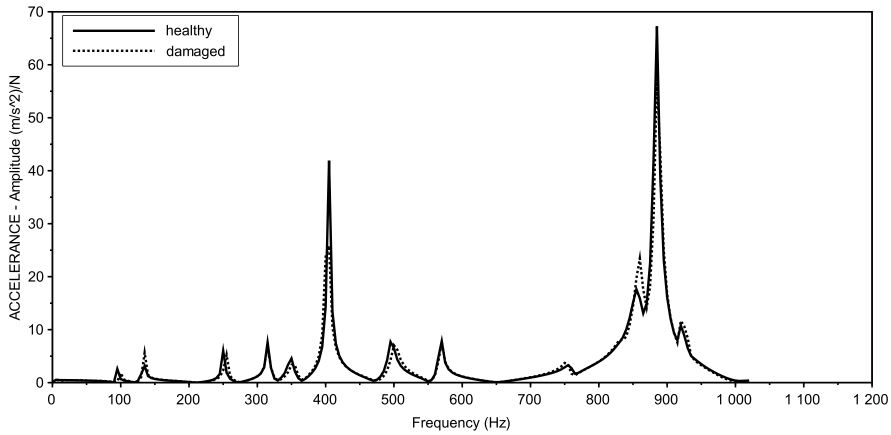

Figure 6. FRFs for healthy and damaged specimen $[0/15/\text{-}15/0/15/\text{-}15]_S$

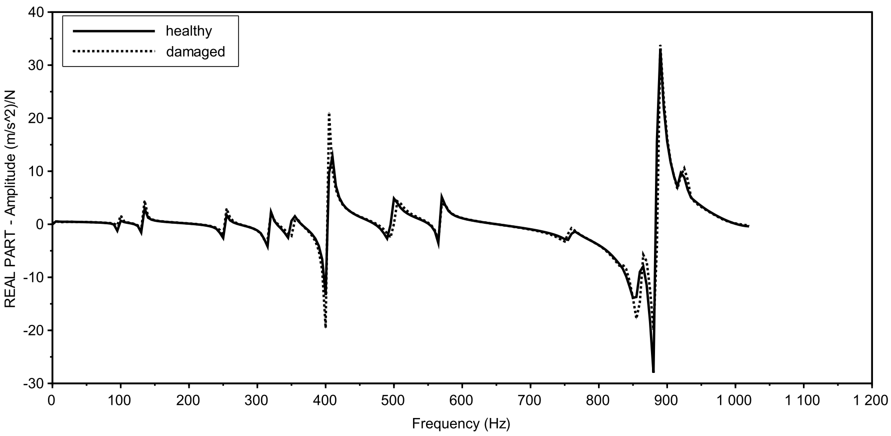

Figure 7. FRFs for healthy and damaged specimen $[0/15/\text{-}15/0/15/\text{-}15]_S$

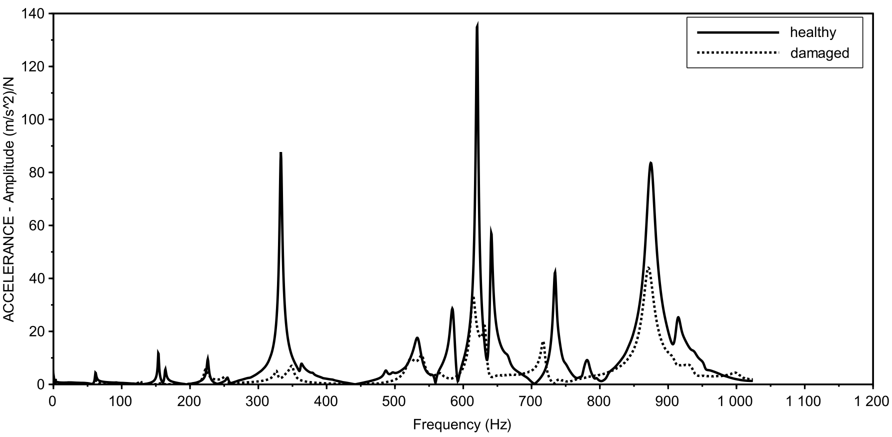

Figure 8. FRFs for healthy and damaged specimen $[0]_8$

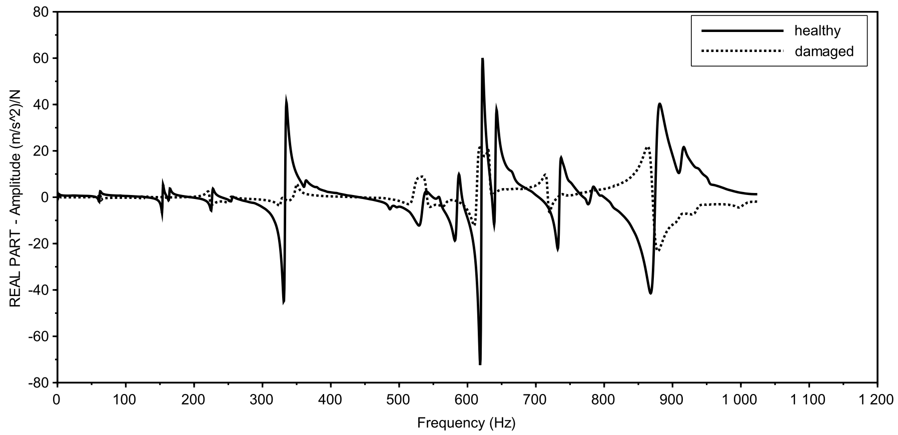

Figure 9. FRFs for healthy and damaged specimen $[0]_8$

Table 3. ANNs topology and parameters: composite plates.

Plates FRFs-cases	Topology	Cost function	Learning rate	Momentum value
$[0]_8$ *(accelerance)*	20-[9,2]-1	L_2	backtracking line search	0.7
$[0]_8$ *(real part)*	20-6-1	L_2	backtracking line search	0.7
[0/15/-15/0/15/-15]$_S$ (accelerance)	20-12-1	L_2	backtracking line search	0.7
[0/15/-15/0/15/-15]$_S$ (real part)	30-[20,10]-1	Cross-entropy	0.1	0.1

The results are summarized in Tab. 4. The ANN generalizes 100% of the patterns correctly using accelerance values and 87.7% using real part values for [0/15/-15/0/15/-15]$_S$ stacking orientation. For $[0]_8$ plates, the testing accuracy is 100% using real part values and 90% using accelerance values. As observed in the previous studies, there is not a problem of underfitting neither overfitting, since training and testing sets got very good accuracies results.

Table 4. ANNs summary results: composite plates.

Plates FRF-cases	Training iterations	Training accuracy (%)	Validation accuracy (%)	Testing accuracy (%)
$[0]_8$ *(accelerance)*	300	100	83.3	90
$[0]_8$ *(real part)*	150	100	100	100
[0/15/-15]$_{2S}$ (accelerance)	150	100	90	100
[0/15/-15]$_{2S}$ (real part)	400	100	80	87.7

4. CONCLUSIONS

Considering the three mechanical systems, the methodology which combines PCA and ANN through the use of the dynamic responses shows good performance when working with two types of class. Both in the aluminum beams, in rolling bearings, and on composite plates, the methodology shows good accuracy in predict damage presence. As observed in the rolling bearing system, the reduction of the principal component does not interfere with the ANN learning process, showing that the PCA can get the main features of the inputs with just 2 PCs. The feature observation is a very important step when working with PCA and ANN. As seen in the composite plates system, the type of FRF curve (accelerance values or real part values) can influence the behavior of the ANN.

The present work evaluates the performance of the new methodology for SHM, combining vibration-based method, PCA, and ANNs. Three different mechanical systems are studied: metallic beams, rolling bearings, and composite plates. The study shows promising results to validate the methodology to detect damage in the varied materials and structures. For an aluminum beams system case, the methodology generalizes 93.55% of the samples correctly, even with a very tight difference between the healthy, 2mm and 4mm crack sizes patterns. For the rolling bearings and the composite plates systems, the methodology generalizes 100% correctly, without underfitting and overfitting problems.

Therefore, these results show that it is possible to identify the damage in components in terms of dynamic behavior. In other words, there is a good perspective for the application of this methodology in Structural Health Monitoring (SHM) systems for structural components.

5. ACKNOWLEDGEMENTS

The authors acknowledge the financial support of the State Research Founding Agency (FAPESC process number: 2017TR1747 and 2017TR784). As well as, São Paulo State Research Foundation (FAPESP process number: 2012/01047-8 and 2015/15221-8), Coordination for the Improvement of the Higher Level Personnel (CAPES process number: 011214/2013-09), National Council for Scientific and Technological Development (CNPq process number: 401170/2014-4 and 310094/2015-1). The authors also would like to thank Navy Technological Centre (CTM – Brazil) for manufacturing specimens. Also, the authors thank Júlio Zago and Guilherme Zago for the previous studies and the data contribution for the rolling bearings.

6. REFERENCES

Bandara, R.P., Chan, T.H. and Thambiratnam, D.P., 2014. "Frequency response function based damage identification using principal component analysis and pattern recognition technique". *Engineering Structures*, Vol. 66, pp. 116–128.

Bishop, C.M., 2006. *Pattern recognition and machine learning*. springer.

Bishop, C.M. *et al.*, 1995. *Neural networks for pattern recognition*. Oxford university press.

Da Silva, I.B.V. and Adeodato, P.J., 2011. "Pca and gaussian noise in mlp neural network training improve generalization in problems with small and unbalanced data sets". In *The 2011 International Joint Conference on Neural Networks*. IEEE, pp. 2664–2669.

De Medeiros, R., 2016. *Development of a criterion for predicting residual strength of composite structures damaged by impact loading*. Ph.D. thesis, Universidade de São Paulo.

Goyal, D. and Pabla, B., 2016. "The vibration monitoring methods and signal processing techniques for structural health monitoring: a review". *Archives of Computational Methods in Engineering*, Vol. 23, No. 4, pp. 585–594.

Gul, M. and Catbas, F.N., 2009. "Statistical pattern recognition for structural health monitoring using time series modeling: Theory and experimental verifications". *Mechanical Systems and Signal Processing*, Vol. 23, No. 7, pp. 2192–2204.

Haykin, S., 2007. *Redes neurais: princípios e prática*. Bookman Editora.

Li, J., Dackermann, U., Xu, Y.L. and Samali, B., 2011. "Damage identification in civil engineering structures utilizing pca-compressed residual frequency response functions and neural network ensembles". *Structural Control and Health Monitoring*, Vol. 18, No. 2, pp. 207–226.

Nielsen, M., 2018. "Improving the way neural networks learn". Neural Networks and Deep Learning <http://neuralnetworksanddeeplearning.com>.

Revels, J., Lubin, M. and Papamarkou, T., 2016. "Forward-mode automatic differentiation in julia". *arXiv preprint arXiv:1607.07892*.

Samali, B., Dackermann, U. and Li, J., 2012. "Location and severity identification of notch-type damage in a two-storey steel framed structure utilising frequency response functions and artificial neural network". *Advances in Structural Engineering*, Vol. 15, No. 5, pp. 743–757.

Simard, P.Y., Steinkraus, D., Platt, J.C. *et al.*, 2003. "Best practices for convolutional neural networks applied to visual document analysis." In *Icdar*. Vol. 3.

Völtz, L.R., Cardoso, E.L. and De Medeiros, R., 2017. "Damage detection on aluminum beams using vibration-based method and artificial neural networks". Proceedings of the 24th International Congress of Mechanical Engineering - COBEM 2017.

Yuan, F.G., 2016. *Structural health monitoring (SHM) in aerospace structures*. Woodhead Publishing.

Zago, G., 2018. "Detecçao de dano estrutural através da análise de frf por meio de redes neurais artificiais auxiliadas pela análise de componentes principais". Universidade do Estado de Santa Catarina.

Zago, J., 2016. "Detecção de falhas em rolamentos através de análise de vibrações e redes neurais artificiais". Universidade do Estado de Santa Catarina.

Zang, C. and Imregun, M., 2001. "Structural damage detection using artificial neural networks and measured frf data reduced via principal component projection". *Journal of sound and vibration*, Vol. 242, No. 5, pp. 813–827.

RESPONSIBILITY NOTICE

The authors are the only responsible for the printed material included in this paper.