Meccanica (2014) 49:305–324
DOI 10.1007/s11012-013-9795-7

Rotor crack identification based on neural networks and modal data

J.L. Zapico-Valle · E. Rodríguez · M. García-Diéguez · J.L. Cortizo

Received: 13 March 2012 / Accepted: 31 July 2013 / Published online: 23 August 2013

Abstract A model-based procedure for rotor crack localization and assessment is presented in this paper. The procedure is applied to a small-size test rig provided with a notch. Both the position and depth of the notch are estimated through a neural network on the basis of the first four natural frequencies of the rotor. A 3-D finite element model is used to generate the data for training the net. One of the contributions of this paper consists of a meshing procedure that reduces the systematic errors of the model, which have a significant influence in identification accuracy. A sensitivity analysis has been carried out for any size and position of the notch, which constitutes another original contribution in this field. In the studied case, the proposed procedure is able to predict both the position and depth of the notch when the notch depth is greater than 20 % of the rotor diameter. The sensitivity analysis reveals that there are blind spots in the rotor as regards notch identification.

J.L. Zapico-Valle (✉) · M. García-Diéguez
Department of Construction and Manufacturing Engineering, University of Oviedo, Campus de Gijón 7.1.16, 33203 Gijón, Spain
e-mail: jzapico@uniovi.es

E. Rodríguez · J.L. Cortizo
Department of Construction and Manufacturing Engineering, University of Oviedo, Campus de Gijón 5.1.22, 33203 Gijón, Spain

Keywords Rotor crack · Crack identification · Finite element model · Neural networks

1 Introduction

Rotating machines are extensively used in many applications such as power generation, marine and aircraft engines, etc. Most of these machines contain horizontal rotors that experience bending normal stresses due to gravitational loads. When rotating, periodic tension-compression cycles are generated. Under these circumstances, rotors are prone to develop transversal fatigue cracks that grow progressively. When the cracks reach a critical size, the rotor breaks and the machine bursts. The consequences of these accidents are usually catastrophic, with deaths and injuries in addition to significant damage to surrounding installations. An introduction to case histories of cracks in industrial machines since the 1950's can be found in ref. [1].

It is obvious that these catastrophic effects should be prevented for safety as well as economic reasons. In fact, extensive studies have been carried out since the early 1970's aimed at developing methods for crack identification in rotors. At the end of last century there was a significant body of research on crack identification, which was reviewed in refs. [2–4]. Recent developments in this field are reported in [5, 6]. The methods to identify rotor cracks are arranged here in five consecutive levels, based on Rytter's original arrangement [7] and the extension provided by Worden et al.

[8] for damage identification. The arrangement has a hierarchical structure; each level requires the completion of all the previous ones. The levels are:

1. Detection: the method gives a qualitative indication that a fault is present in the machine.
2. Classification: the method indicates whether the fault is a rotor crack.
3. Localization: the method provides information about the probable position of the crack in the rotor.
4. Assessment: the method provides information about the shape and size of the crack.
5. Prediction: the method estimates the remaining useful life of the rotor.

Detection methods are essential to prevent the catastrophic effects of rotor fracture and are well established nowadays. They comprise on-line methods based on spectral analysis of the operational vibration, which is continuously or periodically measured. When changes in the $1\times$, $2\times$ and $3\times$ rev. components occur and are statistically significant, the presence of a fault in the machine is suspected. However, these symptoms can be caused by many different faults: unbalance, bow, coupling misalignments, polar asymmetries in the generator, journal ovalization, nonlinear effects in oil film bearings, etc. In practice, it is necessary to ascertain the actual causes of the symptoms so as to repair the machine properly. This classification is accomplished by off-line methods based on vibration originating from run-down transients. The $2\times$ rev. component at half critical speed is a good indicator of the presence of cracks [4]. Another method based on measurements taken during heating transients is reported in ref. [9] as an effective way to classify crack faults. Recently, a method for detection of bearing faults interacting with rotor crack has been published by Kankar et al. [10]. The features used for the identification are extracted from time-domain vibration signals using statistical techniques. The classification is carried out by means of neural networks (NNs) and support vector machines.

Methods for localization, assessment and prediction are important for decision-making purposes in practice. Information about the location and size of a crack allows the remaining useful life of the rotor to be predicted. Once the presence of a crack has been detected, the replacement of the rotor can be postponed until the size of the crack reaches unacceptable levels, thus minimizing the economic implications of the fault.

A variety of methods for localization and assessment have been proposed in literature. However, they are not as well established as those of detection and classification. These methods can be split in two groups depending on whether machines are rotating or not. A number of prominent papers are outlined next beginning with those belonging to the first group. Pennacchi et al. [11] developed a method for crack location based on spectral variations of the components of the rotation speed that are obtained from the measurements at the bearings during run-down transients. It is a model-based method in which the effect of the faults is replaced by an equivalent system of external moments that are identified by a least-squares procedure. The crack depth is calculated by comparing the static bending moment due to the rotor weight with the identified periodical bending moment. The method is validated through a medium-size test rig. Sawicki et al. [12] proposed a method for crack identification likewise based on spectral analysis of the vibration measured at the bearings. Monitoring is done during normal operation employing auxiliary sinusoidal excitation. Identification is based on a simplified finite element model considering a breathing crack. After applying the method on a test rig, these authors conclude that the method is not robust enough to locate and quantify a crack and further development is needed.

Several papers belonging to the second group are outlined next. Xiang et al. [13] developed a model-based crack identification method. The system is modeled using finite element wavelet method. The crack identification problem is solved through the contour-plotting method on the basis of the first three natural frequencies of the rotor system. The method is experimentally verify by a small-scale rotor with a single notch at three different positions in the shaft. They found that an appropriate modelling of the system details and the boundary conditions is essential for reliable predictions. A method based on modal data is presented by Dong et al. [14], who use the first three natural frequencies of the rotor under stationary conditions and a wavelet finite element model. An experimental procedure is proposed to obtain highly precise modal parameters. The procedure is validated by a small-size test rig with a single notch. Yu et al. [15, 16] proposed an inverse identification method based on NNs. The method is tested through a small-scale de Laval rotor. The inputs of the net are the first mode shape ordinates. The data for training the net are obtained from a finite

element model composed of bar elements. The method is tested by numerical simulation, no experimental validation being provided. The authors conclude that a more accurate finite element model is needed.

In this paper we present a model-based method for rotor crack location and assessment. Several innovations are proposed in order to enhance the reliability of the crack identification. We suggest the use of modal data in a frequency range higher than that used in the actual methods as a way of increase the accuracy of the identification. A 3-D finite element model has been used for the first time to model the complete rotor system. This model describes the geometrical details and boundary conditions more adequately than the line models used in the existing methods in literature; the model is thus capable to reproduce more accurately the proposed high modes. Another innovation in this paper is the study of the robustness of the identification for any position of the crack in the shaft, which is only checked for a fixed position of the crack in previous papers. The use of NNs for the crack identification allows the sensitivity to the input error to be easily studied. The paper starts with a justification of the adopted identification procedure. A detailed description of the experiments, NNs and finite element model is then given. Finally, the identification procedure, the sensitivity analysis and the discussion of results are presented before proceeding to the Conclusions section.

2 General approach

The method proposed here is intended for machines under stationary conditions, i.e. when they are not rotating. Once the presence of a crack has been detected by other procedures, the method would be applied during the subsequent shutdown. The method could also be applied during periodic overhands. The method is proposed as an alternative or complement to other localization methods such as visual inspection, dye penetrant inspection, etc.

Impact tests are proposed to obtain vibration data. They can be carried out by providing the machine with adequate access holes. Thus, the machine does not need to be completed dismounted for the test, and standard sensors can be used at the bearings to measure the response.

Natural frequencies of the rotor are selected as features for crack identification because they constitute accurate, informative parameters that can be easily obtained from the adopted tests. The use of high frequencies is proposed as a way of increasing the sensitivity of the procedure. This is argued in [17] as a fundamental axiom of structural health monitoring.

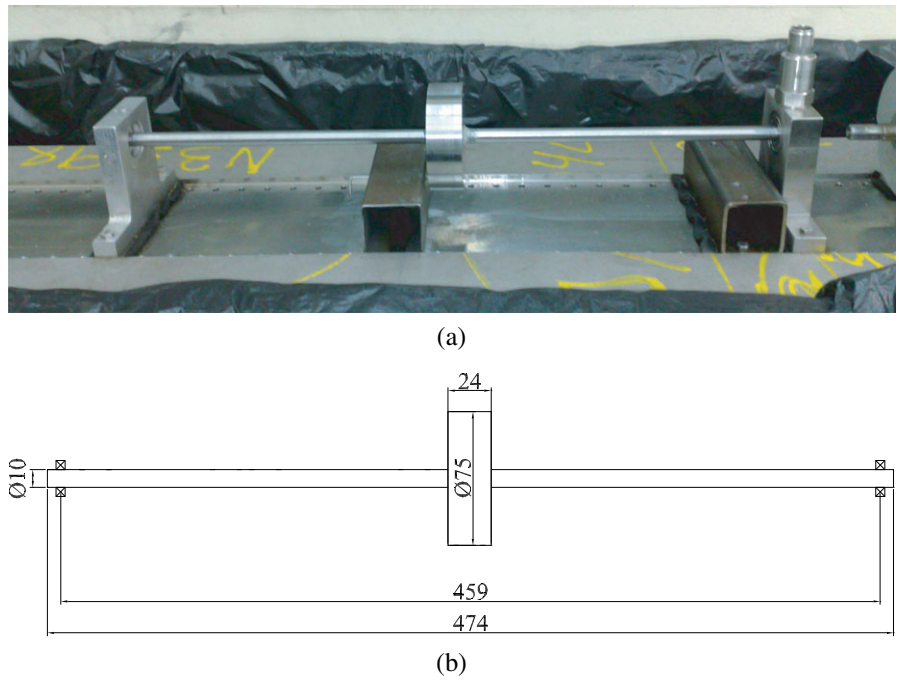
As pointed out in the previous section, line finite element models are appropriate for low frequency vibration, but they cannot adequately reproduce high frequency modes. Hence, a 3-D finite element model is adopted in this study as a way to reduce the model uncertainty even though it represents a higher computational cost. This constitutes one of the main contributions of this paper.

NNs are very suitable for identifying a low number of parameters, as has been verified in previous papers by the authors on structural applications [18–21]. One advantage of NNs is that ill-conditioned data is detected during the training and testing processes because they are unable to generalize in these cases. Thus, improper identification of parameters is avoided beforehand. Another advantage is that, once trained, they are able to provide the output almost instantaneously. This allows a large number of simulations to be carried out in a short period of time, whereas the computation time is almost prohibitive when using direct methods. The influence of the input errors on the predicted parameters can thus be easily performed via Monte Carlo simulations. The study of the robustness of the procedure for any size and position of the crack constitutes another fundamental contribution of this study.

3 Experimental part

3.1 Description of the rotor

A laboratory test rig is used as the experimental reference in this work. It is a small-scale rotor that comprises a shaft plus a disk, both made of steel, and a stiff stainless steel bracket resting on a sand bed so as to isolate it from ground vibration. Two bearing houses of aluminium are bolted to the bracket. The shaft is 10 mm in diameter and 474 mm long and is fitted to two deep grove ball bearings mounted into the bearing houses. The disk is 75 mm in diameter and 24 mm thick and is fixed in the middle of the shaft by means of epoxy resin. The disk contains four longitudinal blind holes uniformly distributed around a circumference so



(a)

(b)

Fig. 1 Initial configuration. (**a**) Photograph. (**b**) Scheme

as to fix its angular position through a pin sliding in a support fixed to the bracket near the disk (see Fig. 1).

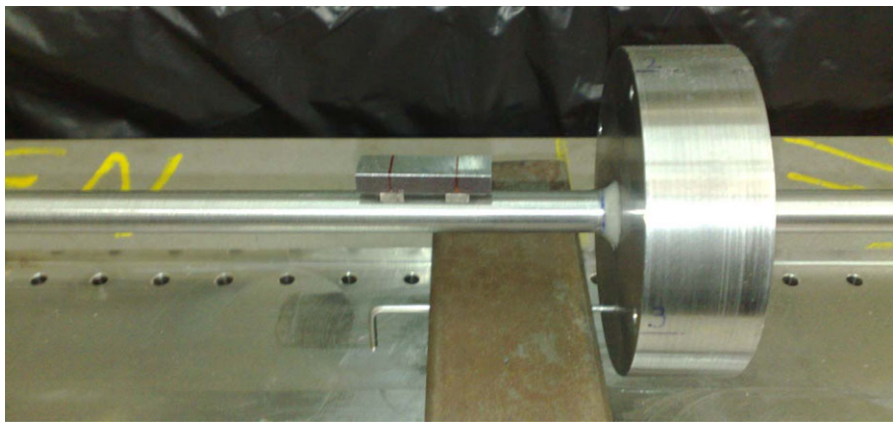
Besides the initial undamaged scenario, two additional sets of configurations were taken into account. The first one corresponds to four mass modifications labeled M-1 to M-4. The second one includes five different levels of damage labeled D-1 to D-5. A movable mass of 11.5 g was used to modify the mass of the rotor. This consists of two small magnets stuck with epoxy resin on a steel bar with a clearance of 15 mm. It was placed at the bottom of the disk (M-1), and at the top of the left semi-shaft close to the disk (M-2), in the middle (M-3) and close to the bearing (M-4) (see Fig. 2). The damage configurations consisted of a horizontal cross notch obtained by sawing the shaft at the same position, p, 84 mm from the left end. The thickness, t, of the notch was approximately 1 mm, while its depth, h, was 1.0, 1.8, 3.1, 3.7 and 4.7 mm respectively for configurations D-1 to D-5 (see Fig. 3).

3.2 Modal testing

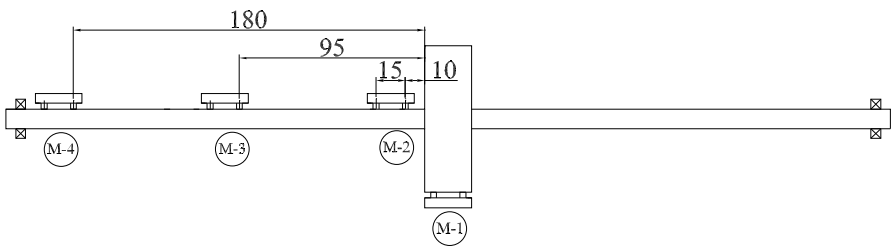
The rotor was equipped with an IMI accelerometer with a 100 mV/g sensibility. It was screwed to the top of the right bearing house in the upright position. The rotor was dynamically excited by means of a small hammer with a mass of 33 g. The impacts were applied in a vertical direction on the top of the disk close to the corner (see Fig. 4). In this position, the first four vertical modes of vibration of the rotor are well excited. CSI machinery analyzer was used to process the signals. This supplies a list of natural frequencies of the rotor based only on the output signal from the accelerometer.

A preliminary set of modal tests was carried out for the four selected angular positions of the rotor. A significant discrepancy in frequencies was found as a function of the angular position. This indicates that the experimental rotor is not exactly axisymmetric. Consequently, it was decided to keep the rotor in the same angular position for the remainder of the experiments so as to avoid the corresponding variability in the results.

A set of 32 experiments was carried out on the initial configuration in order to identify the statistical properties of the results. Table 1 contains the mean value of the identified natural frequencies along with the standard deviation of the set. The scatters of all the frequencies have the same order of magnitude, the standard deviation ranging between 0.27 Hz in the first mode to 1.06 Hz in the third one.

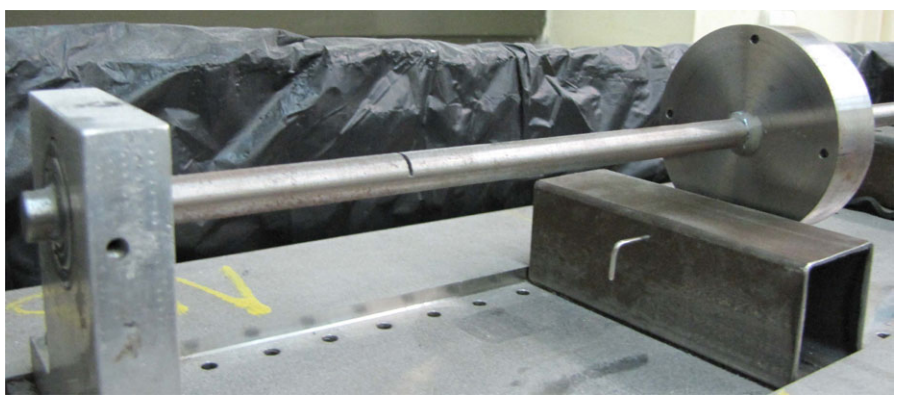


(a)

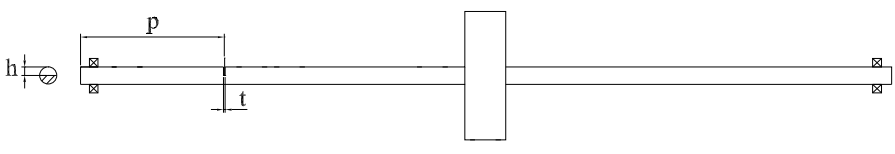


(b)

Fig. 2 Mass configurations. (**a**) Photograph. (**b**) Scheme



(a)



(b)

Fig. 3 Damage configurations. (**a**) Photograph. (**b**) Scheme

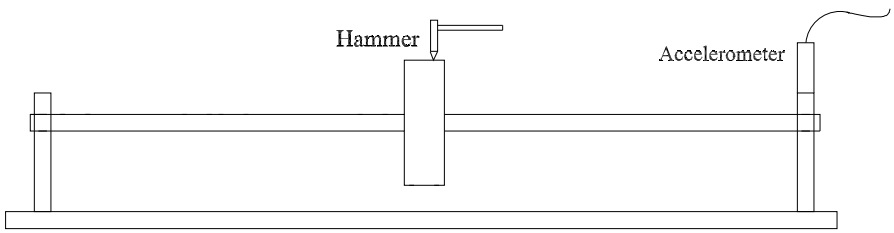


Fig. 4 Scheme of the experimental set up

Table 1 Experimental results. n: number of tests. f: mean natural frequency. (σ): standard deviation

Configuration	n	f_1 [Hz] (σ_1)	f_2 [Hz] (σ_2)	f_3 [Hz] (σ_3)	f_4 [Hz] (σ_4)
Initial	32	41.95 (0.27)	317.65 (0.48)	668.08 (1.06)	901.04 (0.64)
M-1	16	41.68 (0.23)	315.22 (0.45)	667.99 (0.72)	891.47 (0.46)
M-2	16	41.65 (0.34)	316.82 (0.48)	667.71 (1.05)	900.18 (0.78)
M-3	16	41.94 (0.30)	311.65 (0.29)	647.43 (0.57)	887.96 (0.84)
M-4	16	41.95 (0.25)	316.99 (0.34)	661.41 (0.81)	891.69 (0.82)
D-1	16	41.85 (0.32)	317.96 (0.44)	667.99 (0.49)	900.69 (0.78)
D-2	16	41.87 (0.21)	317.34 (0.24)	666.20 (0.61)	898.48 (0.74)
D-3	16	41.63 (0.27)	315.31 (0.41)	660.16 (0.62)	891.52 (0.57)
D-4	16	41.60 (0.21)	314.31 (0.41)	655.72 (0.87)	887.90 (0.73)
D-5	16	41.59 (0.27)	309.76 (0.29)	640.07 (0.65)	874.46 (0.82)

A set of 16 modal tests was carried out for each mass and damage configuration. Table 1 also includes these results. As can be seen, the scatters of these sets have the same order of magnitude as those of the initial configuration, thus providing evidence of the regularity of the measurements.

4 Neural networks

NNs are used in many engineering applications for regression and classification purposes. They are computational models inspired in the structure and operation of the human brain. They consist of a series of connected units called neurons. The stiffness of the connections is represented through numerical values called weights and the knowledge is stored using a set of weights. There are different types of NNs, the most popular being the so-called feed-forward multilayer perceptron (MLP) and radial basis functions. The process of calibrating weights is known as training or learning in literature. The training is based on a set of patterns of training data and can be obtained either numerically or experimentally. There exist several training strategies. They can be broadly classified into two groups: batch and sequential learning. In the batch learning process, the architecture of the net is kept fixed and is determined heuristically. Moreover, the complete training data is used in each step of the training process. The calibration of weights is commonly posed as an optimization issue, in which a quadratic error function that represents the discrepancies between the predictions of the NN corresponding to the input and the target data is minimized with respect of the weights. The minimization is usually carried out in an iterative manner. There are two different stages at each step of the minimization process namely, evaluation of the derivatives of the error function and adjustment of the weights. The back-propagation algorithm is usually used for the first task, while several algorithms are available for the second task: gradient descent, conjugated gradients, quasi-Newton methods, Levenberg-Marquardt algorithm [17]. Huang et al. [22] have proposed an advanced algorithm called extreme learning machine (ELM) for training single-

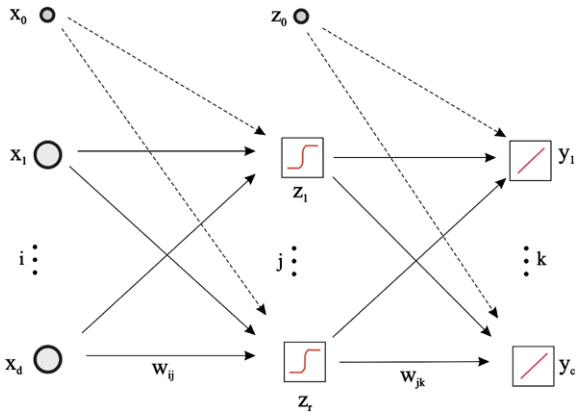


Fig. 5 Configuration of a multilayer perceptron

hidden layer feed-forward neural networks. The input weights and biases are chosen at random, and then the output weights are analytically determined. In general, this algorithm provides better generalization and faster learning speed than the back-propagation algorithms. Suresh et al. [23] developed an extension of the ELM algorithm to measured the visual quality of coded images, which was transformed to a classification problem. They proposed two schemes to select the input weights and bias values such that the generalization performance is maximized. In the sequential learning, the training samples are presented one-by-one or chunk-by-chunk with fixed or varying size and discarded after learning. In addition, the architecture of the net can be refined during the training process. This approach is aimed to maximize the generalization performance and learning speed. Sequential learning is very appropriate for on-line and classification applications. Some related evolving and self-regulating algorithms developed recently can be found in references [24–27].

NNs are used in this paper as an inverse algorithm to solve both the identification of parameters in the FE model updating and the further locating and assessment of damage as a function of the measured modal properties of the rotor. The configuration of NNs used in this work was a two-layer feed-forward MLP. The inputs, x_i, were the natural frequencies of the rotor and the outputs, y_k, the parameters to be identified. Output variables are obtained from the input in a concatenated way, through linear combinations of previous layer values and weights, w, transformed by activation functions. A tangent hyperbolic activation function was used for the hidden layer, g, and a linear function for the output layer, $\widetilde{\mathrm{g}}$ (Fig. 5 and Eq. (1)). Thus, the mathematical relation between input and output can be expressed as:

$$y_k = \widetilde{\mathrm{g}}\left[\sum_{j=0}^{r} w_{kj}^{(2)}\,\mathrm{g}\left(\sum_{i=0}^{d} w_{ji}^{(1)} x_i\right)\right], \tag{1}$$

in which d and r stand for the number of inputs and hidden neurons, respectively.

This constitutes a black-box model based on mother basis functions of a single variable. It can approximate any continuous non-linear multivariate function within a given finite domain by adjusting its weights. The accuracy of the approximation increases with the number of hidden units [17]. The error back propagation algorithm with batch strategy was used for the first step and the scaled conjugate gradients (SCG) algorithm were selected for training the MLPs. SCG is an efficient algorithm of optimization that takes the minimum number of cycles to minimize the error function. SCG can be regarded as a gradient descent algorithm in which the training rate and momentum are automatically optimized at each cycle of training.

The number of hidden neurons needed in each case to reach an appropriate generalization performance is obtained by the following procedure. The available data, which is obtained through a finite element model in this study, is split into two parts. One is used for training, while another is used for testing. Several MLPs are tried in turn increasing the number of hidden neurons. Each configuration is trained using the training data corrupted with low random noise, and the related training error is computed. Afterwards, the testing data is presented to these trained MLPs and from their outputs the errors related to the testing targets are evaluated. If the number of hidden neurons is very low, both the training and testing error are high. The related MLP cannot reproduce accurately the data, it underfits the data. When increasing the number of hidden neurons, the testing error diminishes progressively. However, if the number of hidden units is too large, the MLP tends to fit not only the underlying input-output mapping, but also the noise. This result is called overfitting. When overfitting occurs, the testing errors tends to increase. As a result of this, the testing error diminishes as a function of the number of hidden units until a minimum is reached and then increases. These trends are illustrated in Fig. 6. This figure does not correspond to the cases studied in this

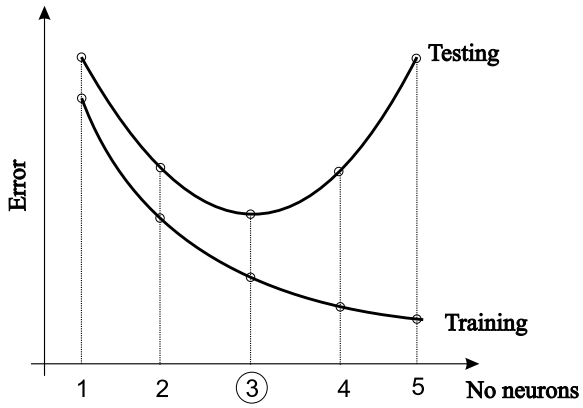


Fig. 6 Testing and training errors versus number of hidden neurons

paper, but is a schematic representation. The configuration corresponding to the minimum training error is related to the proper model order and is chosen for the application. Once established the MLP's architecture, it is eventually trained by using the original data without noise. The NETLAB package was used to compute the MLPs, which is implemented as a set of functions written in the MATLAB language [28] using only core functions. This library is freely offered by its authors from Aston University [29].

5 Finite element model

5.1 Modelling

A 3-D model was developed to reproduce the dynamic behavior of the rotor. It was parametrically established so as to allow modifications to be carried out automatically. The model was coded in ANSYS© [30]. SOLID186 elements were used to model the shaft and the disk. This is a 3-D solid element that exhibits quadratic interpolation functions. It is defined by 20 nodes, each with three degrees of freedom corresponding to translations in the global x, y and z axes. The effect of the bearings and supports on the rotor was modeled by means of COMBIN14 elements. This is a longitudinal tension-compression spring-damper element that has no mass. It is defined by two nodes with three degrees of freedom corresponding to translations in the global x, y and z axes. Only the spring without damper was considered in this application. As to the meshing, two successive approaches were tried. These will be described in the following paragraphs and will be subsequently referred to as irregular and regular meshes.

First, the shaft cross-section was meshed by selecting master nodes uniformly distributed around the circumferential contour and along the vertical and horizontal diameters. The supplementary nodes and the mesh were automatically generated by the code for the complete section from the master nodes. This pattern was repeated in all the considered sections (see Fig. 7). The disk was meshed in a similar way and linked to the corresponding circumferential nodes of the shaft (see Fig. 7). As to the notch, a block with a fixed length was generated as a function of the notch depth. Lines corresponding to the notch edges were generated. In these lines and the corresponding circumferential contours, master nodes were established at a uniform distance approximated to 1.25 mm in each zone. At the ends of the block, the node pattern coincides with that of the shaft. The complementary nodes and the mesh were automatically generated by the code on the basis of these node patterns (see Fig. 7). For a given position of the notch, p, the shaft was split into two parts around the notch block. In each part, the length of the elements was uniformly distributed and approximated to 5 mm. In the section corresponding to the bearings, supplementary nodes were generated in the radial and longitudinal directions and their movement fixed. These nodes were connected to the related ones on the contour of the shaft section through spring elements.

This arrangement is very versatile in the sense that the model can be automatically generated for any position and depth of the notch on the basis of only a few defined parameters. However, several drawbacks were found in its further application.

One drawback is that the resulting mesh is irregular (see Fig. 7). As a consequence, the mode shapes, especially the highest ones, do not coincide with the vertical and horizontal planes. Moreover, a new mesh is generated for each notch configuration. This remeshing gives rise to another inconvenience. The discretization and distortion errors are not constant, but vary significantly as a function of the notch configuration. Analysis of the numerical results reveals that the evolution of the natural frequencies as a function of the notch position exhibits a sharp shape as can be seen in the case shown in Fig. 8. With these rough data it was impossible to train properly the MLP used to identify the damage; it was unable to generalize.

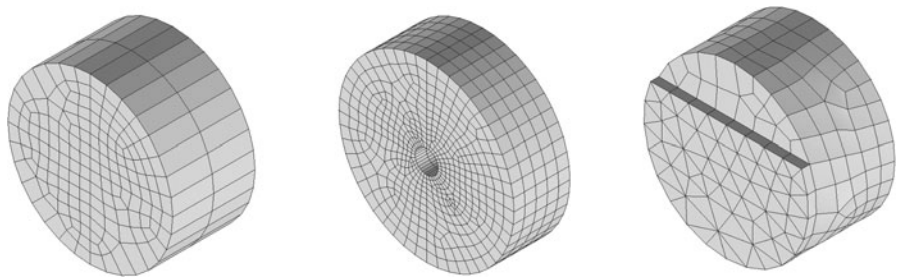


Fig. 7 Irregular meshing. *Left*: shaft. *Middle*: disk. *Right*: notch

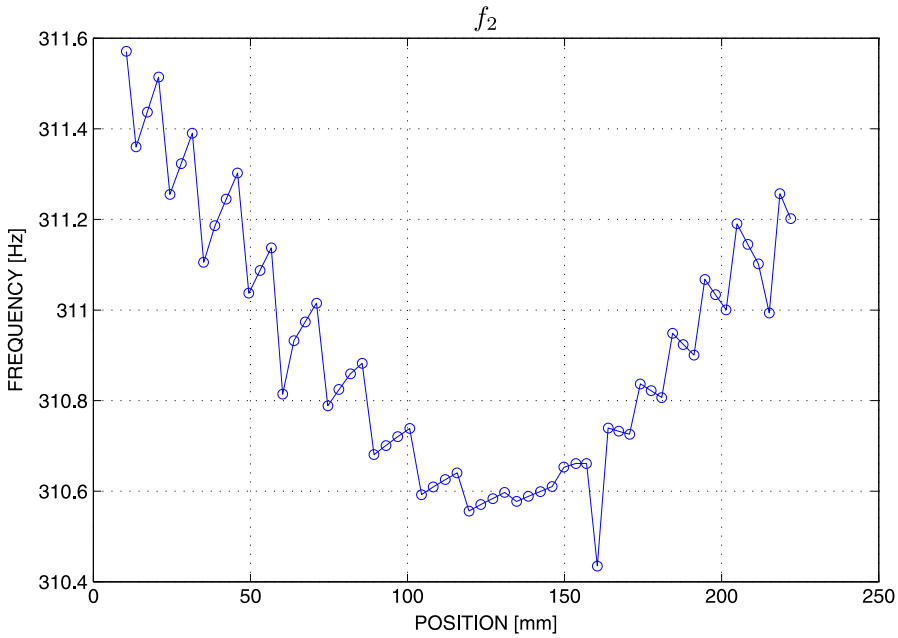


Fig. 8 Evolution of the second frequency as a function of the notch position for $h = 1$ mm and irregular meshing

Eventually, a rather different strategy was adopted for meshing aimed at overcoming the aforementioned drawbacks. Thus, the shaft cross-section was meshed selecting horizontal lines every 0.5 mm. In each line, the nodes were uniformly distributed with a step approximated to 1 mm. The resulting mesh was symmetric with respect to the horizontal and vertical axes (see Fig. 9). Longitudinally, the nodes were uniformly distributed with a fixed step of 2.5 mm. For the disk, a similar regular distribution of nodes was chosen for the longitudinal section. Then, this pattern was uniformly revolved around the disk axis. As a result, an axisymmetric mesh was obtained (see Fig. 9). As to the notch, this was modeled by means of a 10-mm-long block. An identical pattern to that of the shaft was used to mesh the cross-section, the longitudinal step being fixed at 0.5 mm in this case. The notch itself was simulated by simply removing the related finite elements in the block (see Fig. 9). The block was linked by its ends to the other parts of the shaft only at the previously fixed nodes. The part simulating the bearings was meshed in the radial and longitudinal directions in a similar way to that of the irregular meshing (see Fig. 10).

With this arrangement, the notch is simulated for discrete positions and depths. This is not an inconvenience for the damage identification algorithm because it is a regressor that interpolates the numerical data. The advantage of the regular meshing is that the position of the nodes is not altered in the generation of the

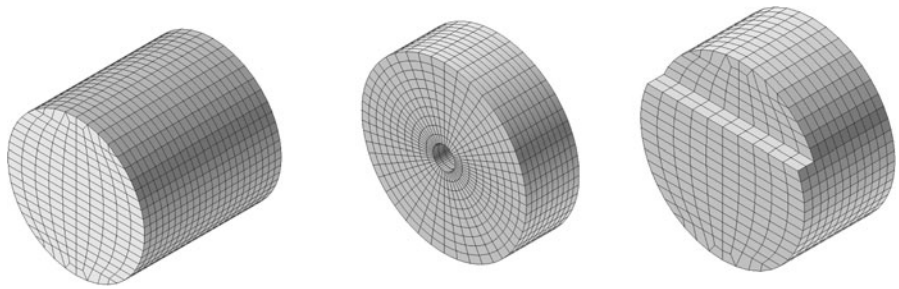


Fig. 9 Regular meshing. *Left*: shaft. *Middle*: disk. *Right*: notch

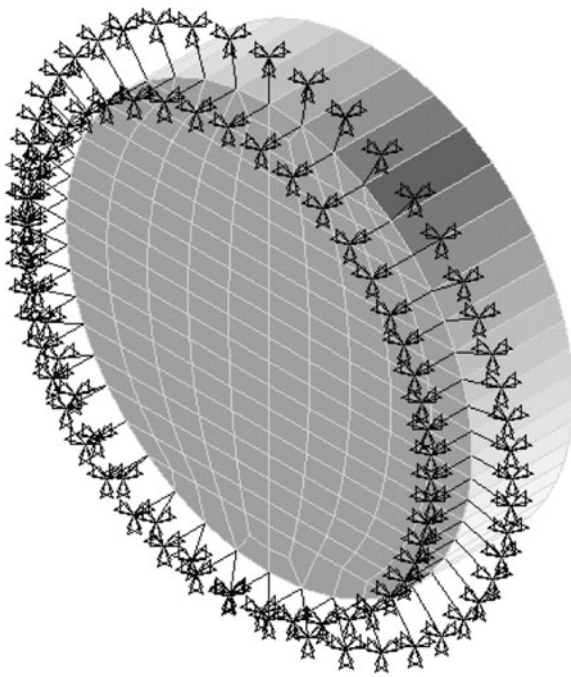


Fig. 10 Bearing modelling

notch, thus avoiding remeshing errors. Thus, an appropriate convergence of the numerical data is hence achieved. This enables proper training and generalization of the damage identification MLP.

5.2 Updating

Model updating can be regarded as a process in which the performance of a model in a given application is enhanced to an acceptable accuracy level. This process is usually based on dynamic tests carried out on the related mechanical system. The results can be expressed in the time, frequency and modal domains. Experiments are conducted at a reduced number of degrees of freedom and limited to a frequency range. In addition, the results always contain measurement and identification errors. Even with these limitations, they constitute the reference to be reproduced as closely as possible by the model.

There are two essential tasks in this process. The first consists in including those features that are missing in the model and are relevant in its dynamic response. The second consists in selecting and updating the numerical values of some model parameters. The dynamic response is calculated from an initial model and comparison with the experimental model allows the missing features to be inferred. These are then added to the model and some relevant parameters are calibrated in such a way that the discrepancies between the analytical and experimental responses are minimized. This updating loop is repeated in successive trials until suitable convergence to the experiments is achieved. The experience and mechanical insight of the analyst, rather than automatic computation, is essential to succeed in both tasks.

5.2.1 Selection of parameters

In this case, the first four vertical natural frequencies of the rotor constitute the available experimental information. In order to avoid an undetermined approach, the number of independent parameters to be updated should be limited to four. The rotor was manufactured with a low dimensional tolerance. Therefore, significant geometrical discrepancies from the design are not expected. Consequently, the geometry of the rotor was set at its nominal shape. As to the density and elastic properties (Young's modulus and Poisson's ratio) of the steel, these values are well known and their variability is very low. Thus, the density was set at its nominal value of 7850 kg/m^3 and kept constant in the

updating process. Young's modulus, E, however, was considered as an updating parameter aimed at correcting not only the effects of the global stiffness errors, but also indirectly those of mass and geometry. If the global mass and stiffness are varied simultaneously, the updating process could become ill-conditioned. This is because the same proportional variation of the global mass and stiffness does not affect the values of the modal variables; different values of the mass and stiffness can give the same value of the natural frequencies.

The boundary conditions are features that have a significant influence in the dynamic response of the model. In this case, they are described in the model by the stiffness of the related horizontal and radial spring elements modelling the bearing and support system. Additionally, it is assumed that both bearings have the same stiffness. Thus, the stiffness of the horizontal and radial springs, k_h and k_r, are chosen as parameters to be updated. In contrast to Young's modulus, prior calibration of the new parameters is not available. Consequently, their updating value should be sought within a broad band including several orders of magnitude.

Eventually, there are three unknown parameters, E, k_h and k_r, and four available natural frequencies. The case, therefore, is overdetermined.

5.2.2 *Updating procedure*

As processing the FE model is time consuming, a strategy aimed at minimizing the computation has been adopted in this study. It consists in finding a first approximation of the parameters by searching within a global space to then obtain the solution through a regressor trained in a local space around the previous approximation.

The updating is posed as an optimization issue. The chosen model parameters are updated in such a way that a given metric of the discrepancies between the numerical and experimental natural frequencies, ε, is minimized. This is defined as a root mean square normalized error

$$\varepsilon = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{f_i^{(a)} - f_i^{(e)}}{f_i^{(e)}}\right)^2}\,100, \tag{2}$$

in which f denotes natural frequency, superscripts (a) and (e) are refer to the numerical and experimental results, respectively, and $N = 4$ is the number of frequencies considered.

Table 2 Values of the parameters for the coarse grid

E [Pa]	k_r [$\frac{\mathrm{N}}{\mathrm{m}}$]	k_h [$\frac{\mathrm{N}}{\mathrm{m}}$]
2.1×10^{11}	5×10^4	1×10^4
2.2×10^{11}	$\mathbf{1 \times 10^5}$	5×10^4
$\mathbf{2.3 \times 10^{11}}$	5×10^5	$\mathbf{1 \times 10^5}$
2.4×10^{11}	1×10^6	5×10^5
2.5×10^{11}	5×10^6	1×10^6

Table 3 Values of the parameters for the fine grid

E [Pa] $(\times 10^{11})$	k_r [$\frac{\mathrm{N}}{\mathrm{m}}$] $(\times 10^5)$	k_h [$\frac{\mathrm{N}}{\mathrm{m}}$] $(\times 10^5)$
2.10	0.50	0.50
2.15	0.75	0.75
2.20	1.00	1.00
2.25	1.25	1.25
2.30	1.50	**1.50**
2.35	**1.75**	1.75
2.40	2.00	2.00

Table 4 Parameter intervals for training and testing the MLP

	E [Pa] $(\times 10^{11})$	k_r [$\frac{\mathrm{N}}{\mathrm{m}}$] $(\times 10^5)$	k_h [$\frac{\mathrm{N}}{\mathrm{m}}$] $(\times 10^5)$
Upper bound	2.20	2.0	1.75
Lower bound	2.10	1.5	1.25

The process has been split into three stages. In the first one, the order of magnitude of the uncertain parameters, k_h and k_r, is sought. The grid search is used for this purpose. The procedure is straightforward. The error function is evaluated for discrete values of the parameters and those corresponding to the minimum error are chosen as an appropriate solution. A coarse grid covering several orders of magnitude was chosen for the parameters k_h and k_r. A narrow interval close to the nominal value was chosen for E because it is a well-known parameter. The search includes all the possible combinations of the chosen values of the parameters, which are shown in Table 2. The solution is highlighted in the same table. As can be seen, the obtained values do not belong to the bounds. This ensures that a minimum of the error function is within the chosen search space. Table 5 includes the natural frequencies and the error corresponding to this initial solution.

Table 5 Results of updating at different stages

	E [Pa] ($\times 10^{11}$)	k_r [$\frac{\mathrm{N}}{\mathrm{m}}$] ($\times 10^5$)	k_h [$\frac{\mathrm{N}}{\mathrm{m}}$] ($\times 10^5$)	f_1 [Hz]	f_2 [Hz]	f_3 [Hz]	f_4 [Hz]	ε [%]
Coarse grid	2.3	1	1	42.21	322.53	663.13	893.59	1.00
Fine grid	2.15	1.75	1.5	41.85	317.49	667.07	899.51	0.17
MLP	2.1449	1.8475	1.5741	41.93	317.60	668.06	900.73	0.03

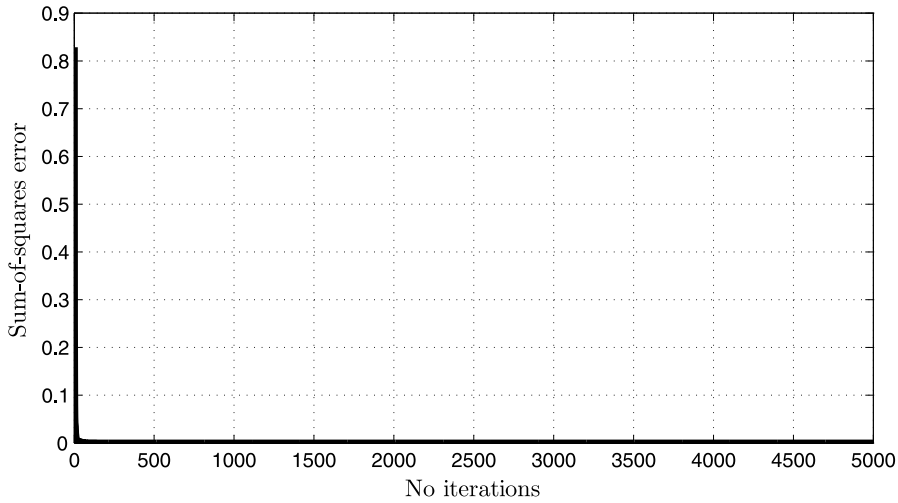


Fig. 11 Updating net. Evolution of the sum-of-squares error as a function of the number of iterations

From this point on, a more precise solution is sought in the second stage by using a fine grid search. The searching values and the results are shown in Tables 3 and 5, respectively. The error is reduced one order of magnitude at this stage.

In the third stage, an inverse procedure based on a MLP is used. The chosen architecture of the MLP is 4:12:3. The inputs are the first four natural frequencies of the rotor, while the outputs are the three updating parameters. The number of hidden neurons was determined by the procedure described in Sect. 4. The data for training and testing the MLP were obtained through the FE model with values of the parameters chosen within narrow intervals around the solution obtained in the previous stage so as to increase the accuracy of the MLP's predictions. The bounds of the intervals are shown in Table 4. Each parameter interval was split into ten uniform subintervals for training purposes and a point was chosen at random within each subspace. A database containing 1000 data sets was thus obtained from the FE model. For testing purposes, 500 points were chosen at random within the complete space defined in Table 4. The evolution of the sum-of-squares error as a function of the number of training cycles is shown in Fig. 11. As can be seen, the errors diminishes rapidly, the final error being very low: 2.378×10^{-5}. The training lasted around 37 minutes in this case. Figure 12 shows the results of testing. The correlation coefficient of the predictions of the trained net with respect to the targets is almost equal to one in all the cases. These results reveal that the net has been properly designed and trained and supplies very accurate predictions.

5.2.3 Results and discussion

After training and testing, the solution was found by feeding the MLP with the experimental natural frequencies. The solution is shown in Table 5. As can be seen, the fitting error is now one order of magnitude lower than the previous one.

A study of the transmission of errors was carried out to ascertain the effect of the uncertainties of the natural frequencies on the results. This was numerically simulated in the trained MLP via the Monte Carlo method. As the natural frequencies are mean

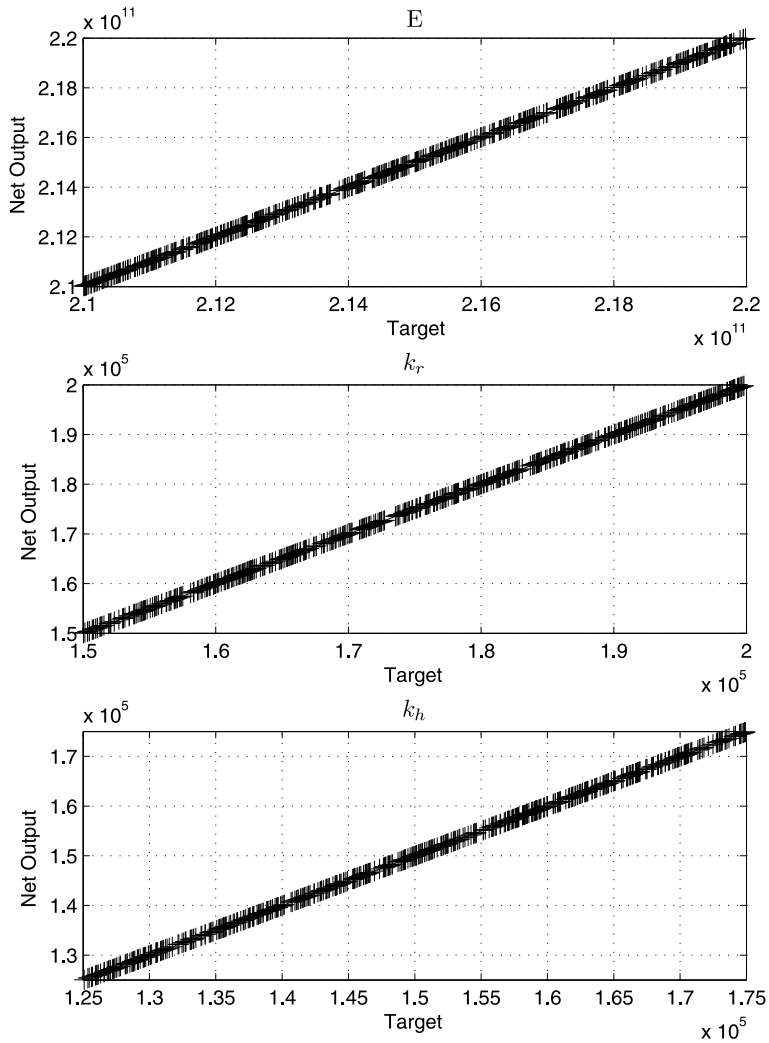


Fig. 12 Updating net. Results of testing

values of a set of 32 measurements, a normal distribution can be considered for this random variable in accordance with the central limit theorem. Thus, each natural frequency is sampled from a Gaussian distribution with mean f_i and standard deviation $\frac{\sigma_i}{\sqrt{32}}$ in the simulation. A database with 100,000 sets was generated and used to feed the MLP and obtain the related solutions. The 95 % confidence interval was achieved selecting the corresponding 97,500 and 2,500 order statistics. Results are shown in Table 6. The intervals are very narrow, especially for the parameter E, which is quite close to its nominal value of 2.07×10^{11} Pa. It should be pointed out that the four natural frequencies provide a good fit just calibrating only three parameters. All these results demonstrate that the updated FE model is suitable to reproduce the given experiments.

Table 6 95 % confidence interval for the obtained parameters

Fractile	E [Pa] ($\times 10^{11}$)	k_r [$\frac{\mathrm{N}}{\mathrm{m}}$] ($\times 10^5$)	k_h [$\frac{\mathrm{N}}{\mathrm{m}}$] ($\times 10^5$)
97.5 %	2.1516	1.9069	1.6564
2.5%	2.1383	1.7915	1.4914

5.2.4 Validation

The updated model is not intended to merely mimic the related experiments, but to reassemble the most significant physical features. Thus, the model should

Table 7 Results of the mass modifications

Configuration	Natural frequencies [Hz]		Discrepancy [%]
	Experimental	FE model	
M-1	41.68	41.68	0.00
	315.22	314.81	−0.13
	667.99	667.77	−0.03
	891.47	891.11	−0.04
M-2	41.65	41.69	0.09
	316.82	315.97	−0.27
	667.71	667.93	0.03
	900.18	898.06	−0.24
M-3	41.94	41.80	−0.33
	311.65	309.84	−0.58
	647.43	646.09	−0.21
	887.96	891.12	0.36
M-4	41.95	41.92	−0.09
	316.99	316.16	−0.26
	661.41	657.11	−0.65
	891.69	887.55	−0.46

be able to reproduce experiments under other loading conditions or after some mass or stiffness modifications. In order to check such a generalization capacity, the experimental mass modifications were simulated in the updated model. The additional masses were modeled as lumped masses at the corresponding nodes of the model. Results are shown in Table 7 along with the experimental data. The analytical frequencies tend to be lower than their experimental counterparts. A possible explanation of this trend could be that the masses add some stiffness to the rotor due to the action of the magnets. Nevertheless, simulations are quite close to the experiments, with discrepancies up to 0.65 %, though most are below 0.5 %. The results validate the capacity of generalization of the model to a certain extent.

6 Damage identification

6.1 Identification procedure

An inverse procedure based on a MLP similar to that of the third stage of the updating was chosen to identify the damage. The first four natural frequencies of the rotor constitute the available experimental information on which the identification will be based. Assuming the damage scenario is a cross notch, this is defined by its position, p, with respect to the left end and its depth, h. Bearing in mind the symmetry of the rotor, only the left half was considered in the analysis.

The data for training the MLP was generated through the updated model. A 1-mm-thick notch was simulated varying its position from 20 to 215 mm every 5 mm. Five different depths, $h = 1, 2, 3, 4$ and 5 mm, were considered in each position and the corresponding natural frequencies were calculated. A database with 200 data sets was thus obtained. Results are depicted in Fig. 13 with the same range for the decrement of all the frequencies. It can easily be seen that the higher the frequency, the higher the sensitivity to the damage, the first frequency, f_1, being almost insensitive to the damage.

Considering in addition that the order of magnitude of the error is similar for all the frequencies, it was decided not to consider the first natural frequency for damage identification. Otherwise, it would add more confusion than information because the magnitude of its inherent error is higher than the frequency change due to damage in most cases. Thus, the MLP configuration adopted in this case was 3:30:2. The inputs are the frequencies, f_2, f_3 and f_4, the outputs are the position, p, and the depth, h, of the notch, the hidden layer containing 30 neurons that were determined through the procedure described in Sect. 4.

The available database was split at random in two subsets containing 140 and 60 data sets. The former was used for training and the latter for testing. Results training and testing are shown in Figs. 14 and 15. As in the updating net, the training process converges rapidly, the final error being: 8.441×10^{-5}. The training lasted around 20 minutes in this case. As for the testing, the predictions of the trained net are excellent.

6.2 Comparison with experimental results

The damage predictions are obtained by feeding the trained MLP with the experimental natural frequencies. Results are shown in Table 8. The influence of the frequency errors on the predictions were evaluated via Monte Carlo simulation. For this purpose, a Gaussian distribution with mean f_i and standard deviation $\frac{\sigma_i}{\sqrt{16}}$ was considered in each frequency. A database with $10,000$ data sets was generated and the corresponding predictions of the MLP were computed. Tables 9 and 10 include the 95 % confidence interval for the position and depth predictions, respectively.

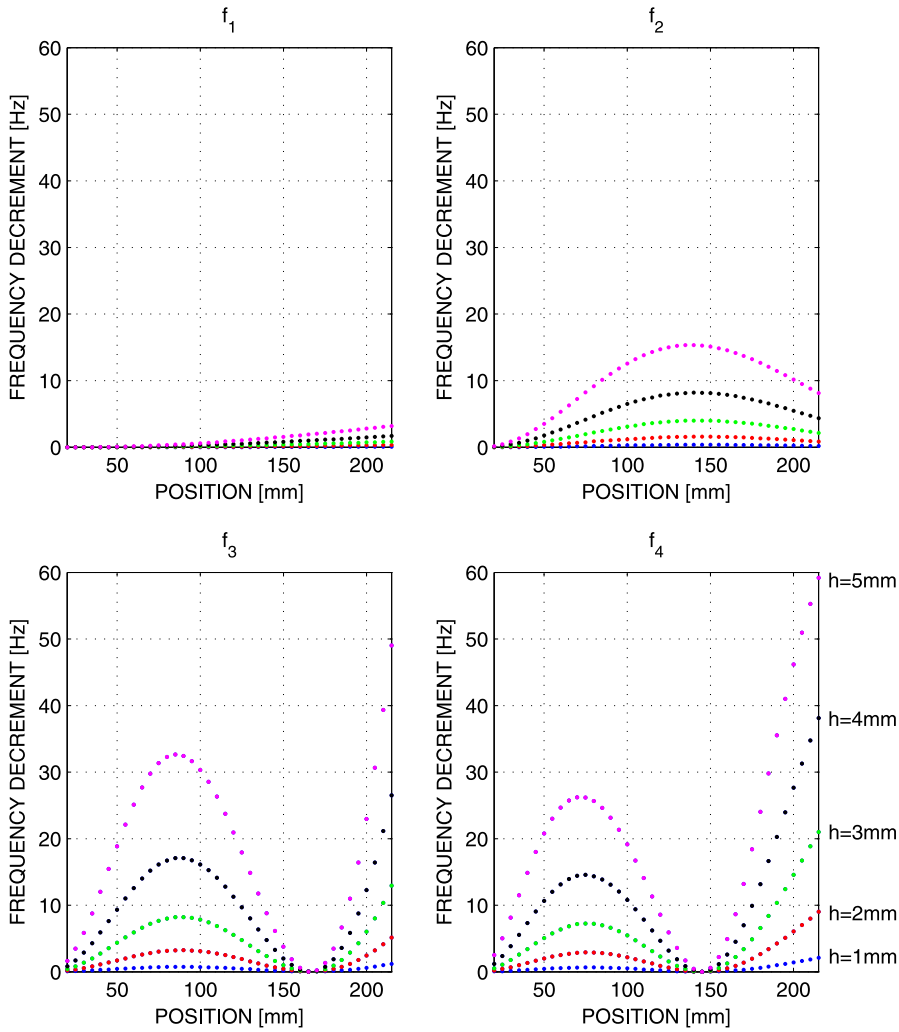


Fig. 13 Frequency sensitivity

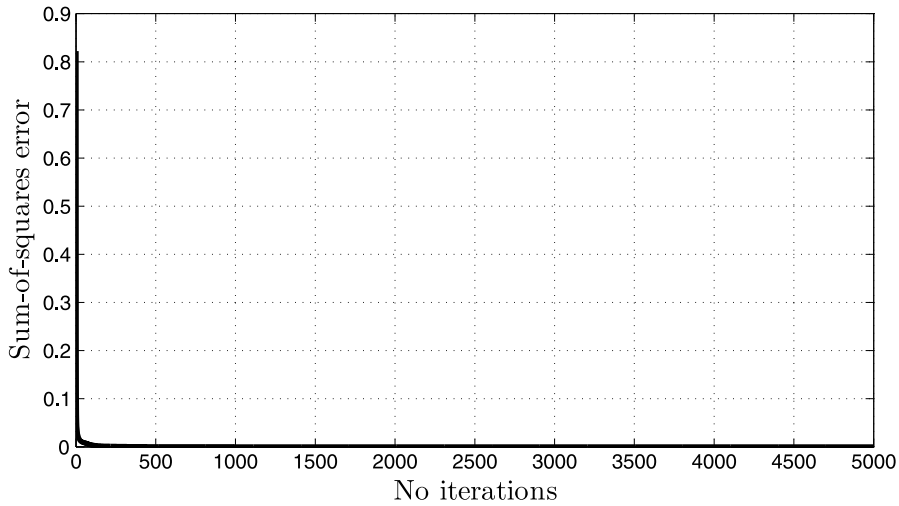


Fig. 14 Identification net. Evolution of the sum-of-squares error as a function of the number of iterations

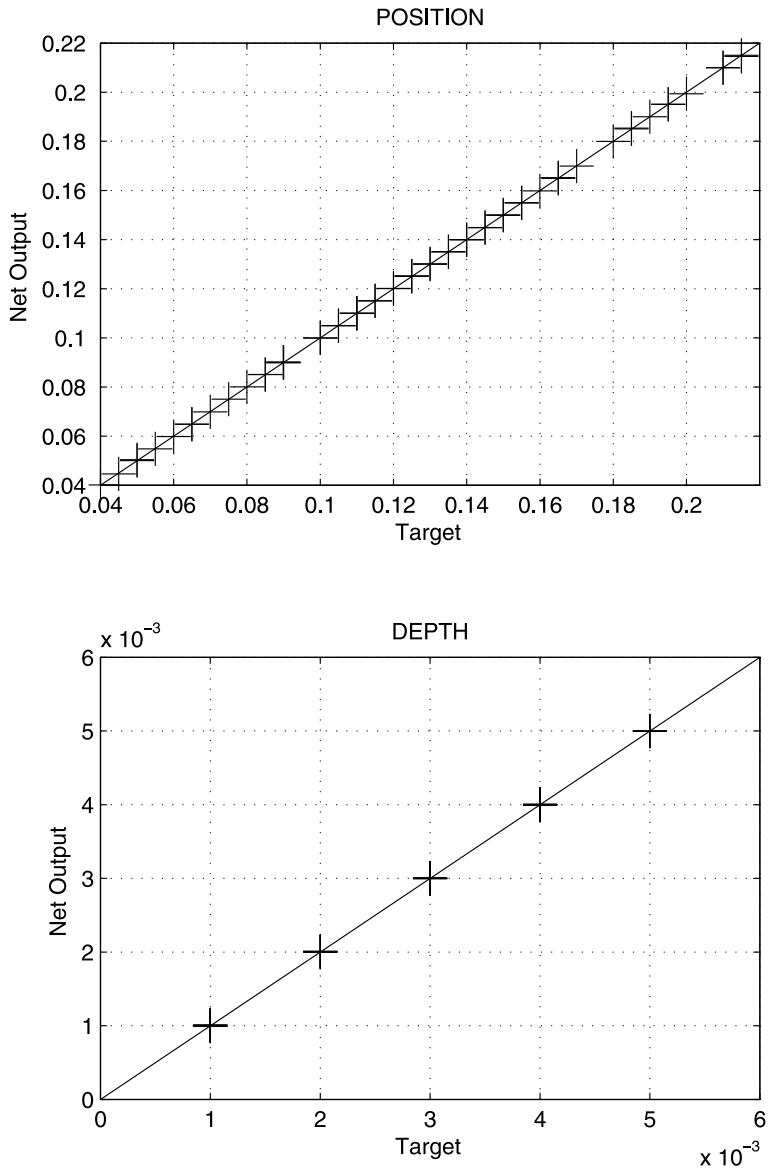


Fig. 15 Identification net. Results of testing

Table 8 Results of the damage identification

Configuration	p [mm]		h [mm]	
	Experimental	Prediction	Experimental	Prediction
D-1	84	−15.26	1.0	27.8
D-2	84	−242	1.8	7.02
D-3	84	86	3.1	3.72
D-4	84	76	3.7	3.91
D-5	84	78	4.7	5.09

Table 9 95 % Confidence interval for the notch position predictions, p [mm]

Configuration	Fractile		Experimental
	2.5 %	97.5 %	
D-1	−1739.4	−791.6	84
D-2	−1960.6	175.7	84
D-3	77.7	93.7	84
D-4	71.0	81.2	84
D-5	75.7	80.0	84

Table 10 95% Confidence interval for the notch depth predictions, h [mm]

Configuration	Fractile		Experimental
	2.5 %	97.5 %	
D-1	18.6	31.5	1.0
D-2	1.2	34.4	1.8
D-3	3.6	3.9	3.1
D-4	3.9	4.0	3.7
D-5	5.0	5.2	4.7

In configuration D-1, the confidence intervals do not include the targets, the position or the depth. Moreover, the confidence intervals are far from the experimental results. This is due to the fact that the decrements in the frequencies are very low at this damage level, even lower than the measurement errors. Thus, the data supplied to the MLP can fall outside the training interval, so the MLP is extrapolating rather than interpolating. Under these conditions, the extrapolation error, which can be substantial, is added to the prediction. The target values of the damage are within the confidence interval in configuration D-2, but the bounds are very large. The previous explanation can be argued in this case. The notch position is within the confidence interval in configuration D-3, and outside but very close to it in configurations D-4 and D-5. As to the notch depth, its target value is outside and slightly smaller than the lower limit in configurations D-3 to D-4.

Summing up, for the studied position of the notch the predictions of the proposed procedure are acceptable when the notch depth is greater than 2 mm, which represents 20 % of the shaft diameter.

6.3 Sensitivity analysis

In the previous subsection, the prediction capacity of the proposed procedure was experimentally studied for a fixed position of the notch. Results cannot be generalized because the predictions also depend on the notch position, p. Therefore, this influence should be studied in order to ascertain the capability of the procedure. An experimental study of the influence of the position is prohibitive because damage is not reversible and a new rotor would have to be manufactured, tested and updated for each position. It was accordingly decided to use the numerical simulation instead.

To this end, the positions and depths used for the training database (see Sect. 6.1) were adopted. For each position and depth, 10,000 samples were obtained from Gaussian distributions with mean f_i and standard deviation $\frac{\sigma_i}{\sqrt{16}}$, and the corresponding damage predictions were obtained through the MLP. The 97.5 % and 2.5 % fractiles are shown in Figs. 16 and 17 for the position and the depth predictions, respectively. These figures are very illustrative of the overall performance of the identification procedure.

Both bounds of the confidence interval are far from the related targets for the lowest notch depth, $h = 1$ mm. This means that this level of damage cannot be identified at any position. For $h = 2$ mm, the predictions are acceptable for values of p from 100 to 175. The damage, however, cannot be identified in the zones close to the bearing and the disk. This provides evidence of the low sensitivity of these zones and the aforementioned extrapolation errors. From this level on, the same trend exists in the predictions. The higher the value of h, the narrower the confidence interval and the lower the length of the blind zones.

The predictions of the proposed procedure for low values of the notch depth and at the blind spots found near the bearing and the disk could be significantly improved by including higher natural frequencies, which are much more sensitive to damage, to the input and reducing their uncertainty. This can be achieved in practice choosing a more appropriate excitation for high frequencies, more accurate sensors and procedures for modal extraction than those used in this study.

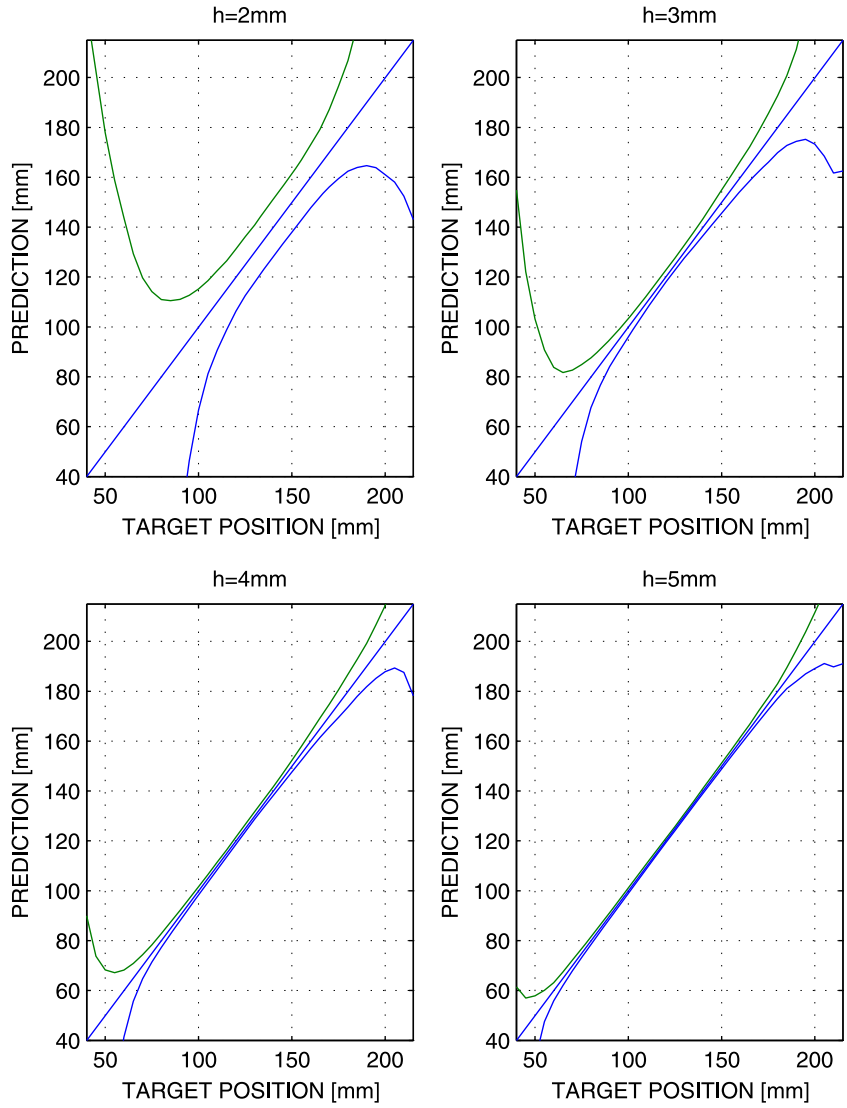


Fig. 16 95 % confidence interval for the notch position predictions

7 Conclusions

A procedure for rotor crack location and assessment that seeks robustness and applicability has been developed. Based on the first four natural frequencies, the procedure was applied to a small-scale test rig provided with a notch with different sizes.

In this particular case, it was found that the procedure is able to estimate both the position and depth of the notch quite accurately when the depth of the notch to diameter of the shaft ratio is higher than 20 %. The procedure becomes unstable for lower values of this ratio. These results could be enhanced by using additional high modes and more accurate modal tests and modal identification.

It has been shown that notch modelling in its standard automatic remeshing version introduces significant systematic errors, which does not allow the generalization of the identification algorithm. The alternative procedure based on the fixed node position overcomes this problem completely. Additionally, a suitable modelling of the boundary conditions is essential to obtain accurate predictions.

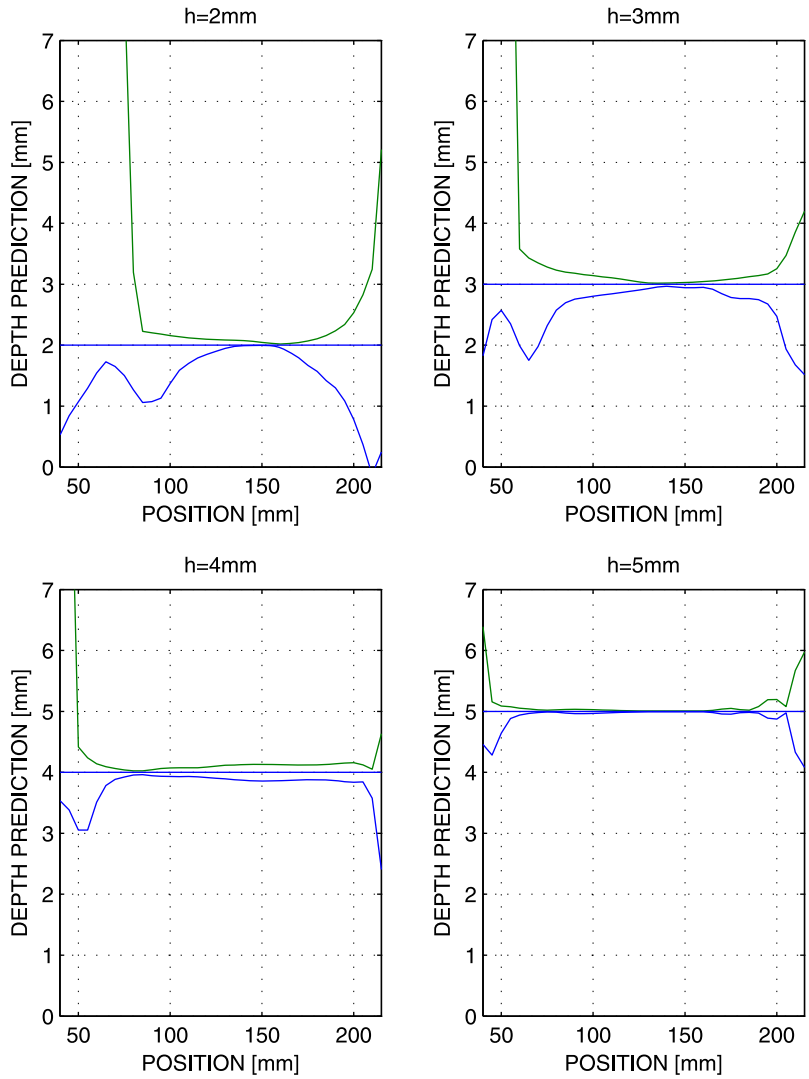


Fig. 17 95 % confidence interval for the notch depth predictions

The sensibility analysis reveals that there exist blind spots, which are close to the bearings and the disk, where notch identification is difficult even for large values of the depth ratio. These results confirm the statement that the robustness check cannot be reduced to a single position, but should be extended to any position on the rotor.

The experience can be extended to cracks by simply introducing special cracked finite elements at the crack tip. The crack breathing effect would be made evident by testing the rotor in different angular positions. The results of this testing would also serve to infer the shape and angular position of the crack.

Acknowledgements Financial support provided by the Spanish Ministry of Education and Innovation through project BIA2006-15266-C02-01 is gratefully appreciated.

References

1. Ishida Y (2008) Cracked rotors: industrial machine case histories and nonlinear effects shown by simple Jeffcott rotor. Mech Syst Signal Process 22:805–817
2. Wauer J (1990) On the dynamics of cracked rotors: a literature survey. Appl Mech Rev 43:13–17
3. Gasch R (1993) A survey of the dynamic behavior of a simple rotating shaft with a transverse crack. J Sound Vib 160:313–332

4. Dimarogonas AD (1996) Vibration of cracked structures: a state of the art review. Eng Fract Mech 55:831–857
5. Bachschmid N, Pennacchi P (2009) Crack effects in rotordynamics. Mech Syst Signal Process 23:1884–1893
6. Lees AW, Sinha JK, Friswell MI (2008) Model-based identification of rotating machines. Mech Syst Signal Process 22:761–762
7. Rytter A (1993) Vibration based inspection of civil engineering structures. PhD Thesis, Department of Building Technology and Structural Engineering, University of Aalborg, Denmark
8. Worden K, Dulieu-Barton JM (2004) An overview of intelligent fault detection in systems and structures. Struct Control Health Monit 3:85–98
9. Bachschmid N, Pennacchi P, Tanzi E, Vania A (2000) Identification of transverse crack position and depth in rotor systems. Meccanica 35:563–582
10. Kankar PK, Sharma CS, Harsha SP (2012) Vibration-based fault diagnosis of a rotor bearing system using artificial neural network and support vector machine. Int J Model Identif Control 15:185–198
11. Pennacchi P, Bachschmid N, Vania A (2006) A model-based identification method of transverse cracks in rotating shafts suitable for industrial machines. Mech Syst Signal Process 20:2112–2147
12. Sawicki JT, Friswell MI, Kulesza Z, Wroblewski A, Lekki JD (2011) Detecting cracked rotors using auxiliary harmonic excitation. J Sound Vib 330:1365–1381
13. Xiang J, Chen X, Mo Q, He Z (2007) Identification of crack in a rotor system based on wavelet finite element method. Finite Elem Anal Des 13:1068–1081
14. Dong HB, Chen XF, Li B, Qi KY, He ZJ (2009) Rotor crack detection based on high-precision modal parameter identification method and wavelet finite element model. Mech Syst Signal Process 23:869–883
15. Yu T, Han Q-K, Qin Z-Y, Wen B-C (2006) Identification of crack location and depth in rotating machinery based on artificial neural network. Advances in neural networks. Lecture notes in computer science, vol 3973, pp 982–990
16. Yu T, Yang Y, Han Q, Yao H, Wen B (2007) ANN-based crack identification in rotor system with multi-crack in shaft. Key engineering materials, vols 353–358, pp 2463–2466
17. Bishop CM (1998) Neural networks for pattern recognition. Oxford University Press, London
18. Zapico JL, González MP, Worden K (2002) Damage assessment using neural networks. Mech Syst Signal Process 17:119–125
19. Zapico JL, Worden K, Molina FJ (2001) Vibration-based damage assessment in steel frames using neural networks. Smart Mater Struct 10:1–7
20. González MP, Zapico JL (2008) Seismic damage identification in buildings using neural networks and modal data. Comput Struct 86:416–426
21. Zapico JL, González-Buelga A, González MP, Alonso R (2008) Finite element model updating of a small steel frame using neural networks. Smart Mater Struct 17
22. Huang G-B, Zhou H, Ding X, Zhang R (2012) Extreme learning machine for regression and multiclass classification. IEEE Trans Syst Man Cybern, Part B, Cybern 42(2):513–529
23. Suresh S, Babu RV, Kim HJ (2009) No-reference image quality assessment using modified extreme learning machine classifier. Appl Soft Comput 9(2):541–552
24. Suresh S, Sundararajan N, Saratchandran P (2008) Sequential multi-category classifier using radial basis function networks. Neurocomputing 71(7–9):1345–1358
25. Huang G-B, Saratchandran P, Sundararajan N (2004) An efficient sequential learning algorithm for growing and pruning RBF (GAP-RBF) networks. IEEE Trans Syst Man Cybern, Part B, Cybern 34(6):2284–2292
26. Babu GS, Suresh S (2012) Meta-cognitive neural network for classification problems in a sequential learning framework. Neurocomputing 81:86–96
27. Babu GS, Suresh S (2013) Meta-cognitive RBF network and its projection based learning algorithm for classification problems. Appl Soft Comput 13(1):654–666
28. Users manual version 6.5 MATLAB (2002). The Math Works, Inc.
29. Bishop CM, Nabney IT (1997) Netlab neural network software. http://www.ncrg.aston.ac.uk
30. ANSYS ® Mechanical vol 11.0. Documentation for ANSYS, ANSYS, Inc.