Detection of various failure causes in complex mechanical systems by the use of Artificial Neural Networks

R. Faglia, M. Tiboni, M. Antonini
University of Brescia, Mechanical Engineering Department, via Branze 38, Brescia
fax +39 030 37024480 - e-mail mtiboni@ing.unibs.it

Abstract The paper presents a methodology based on Artificial Neural Networks (ANN) to perform on-line a diagnosis of the health state of a machinery. The procedure at issue permits to detect the presence of backlash and to determine possible structural failures inside a mechanical system. Backlash and damages are important causes of vibrations in machines, therefore vibrations monitoring gives indirect information on these parasite effects. An ANN is used to classify the system behaviour among a predefined number of classes, receiving as input vibrational signals (simulated or measured). An application is discussed for devices purposely built for indexing motion, where compliance plays an important role affecting the dynamic behavior of the whole machine. An analysis of parameters sensibility for the proposed procedure on simulated cases highlighted the best values and choices for these parameters. Tests of the procedure on experimental data collected on actual devices match closely the good results achieved with simulations.
Keywords : Mechanical Indexing Systems, Elasto-Dynamic Models, Vibrations, Neural Network, Identification.

1 Introduction

The need to improve the standards of quality, to increase efficiency and to reduce production costs has led to an increment of the attention dedicated to diagnostic problems of mechanical systems.

In the past maintenance policy was based on repair after failure; users maintain machines only when something breaks with consequent costs elevation, productivity reduction and, sometime, catastrophic consequences.

In recent years operators and companies have realized that a good maintenance policy "is truly the single largest controllable cost in the operation of a plant or machine" [19]. Different maintenance strategies were consequently developed: preventive, predictive and proactive maintenance.

Preventive maintenance includes activities (i.e. regular care of such components and machine systems, periodic inspections and actions to repair or replace components in impending failure) to prevent impending failure.

Predictive maintenance, also known as condition monitoring, is based on the monitoring of one or more conditions (visual, compliance, vibration and noise, wear debris and heat monitoring) to determine whether material degradation is occurring.

The more recent approach of maintenance is proactive maintenance, an activity performed to detect and correct root cause aberrations of failure.

Several works in literature deal with the development of methodologies to implement condition monitoring maintenance [8, 7, 9], some of them used an Artificial Neural Network for an automatic identification of the system state by giving to the network the current value of the monitored condition. ANN can be successfully used to perform non linear systems identification, while more traditional algorithm for classification failed.

The application fields of these methodologies are manifold: paper-making industry, economy, railway systems, robotic manipulators, rotating machinery, civil engineering, medical diagnosis, etc..

The proposed work deals with the development of a predictive maintenance methodology based on vibrations monitoring, where a Neural Network is employed to identify the faulty operation and the damage level of the condition.

Vibrations were chosen as condition to monitor due to the nature and to the type of the mechanical system considered, a mechanical indexing system, formed by a rotating table driven by an electrical motor through a gear speed reducer.

The existence of an effective link between machine vibrations and its health conditions has been verified several times. Previous studies of the authors [1, 4] demonstrated that backlash in mechanical components, particularly in the speed reducer, and the characteristic of the compliance, inside junctions of mathematical models of the system, play an important role in the amplification of vibrations.

The development of systems based on vibrations measurement as mean to detect the condition of a machine, has been encouraged from the easy application of such technology. Simply a common accelerometer and the respective instrumentation are needed.

The collected acceleration data are usually preprocessed to better identify the type of damage in the system. Data preprocessing techniques go from the analysis of the Fast Fourier Transform (FFT), to the Power Spectral Density (PSD) or to higher order statistics quantities (HOS) [?]. Moreover, identification of breakdown is also carried out through analysis with Wavelet [7] or with algorithms like Dynamic Time Warping (DTW).

The interpretation of results obtained with such analyses can be executed by an engineer, but it is difficult to implement in traditional systems for faults automatic acknowledgment. This derives from the limits connected to the structure of the algorithms that come from the Touring's machine, which are rather efficient in classifying well known cases, but are not capable to generalize situations never seen before or to continue learning even after the completion of the code writing phase.

The great variety of situations likely in the mechanical systems diagnostic, combined to the casuality of such phenomena, leads to the research of new methods for automatic inspection of machine health.

Algorithm based on parallel calculation instead of sequential one, like artificial neural networks, could be effectively used to develop automatic diagnostic tools in the presence of highly non-linear phenomena.

In [2] Lucifredi and others showed that structural failures could be associated with bilinear behaviour, thus a bilinear characteristic in a joint of a model of a mechanical system could be used to simulate a crack in the device.

The proposed procedure of diagnosis has been firstly calibrated and tested on simulated data. Simulations were executed with elasto-dynamic mathematical models (arranged by the authors in previous researches), for which structural failures and backlash were enclosed in opportune characteristics of the joints of the models.

An analysis of parameter sensitivity has been developed to adequately calibrate the procedure.

Experiments executed in a second stage of the research permitted to collect a database of measured acceleration. An actual mechanical system for indexing motion was considered and structural failure and backlash in this case were actually generated.

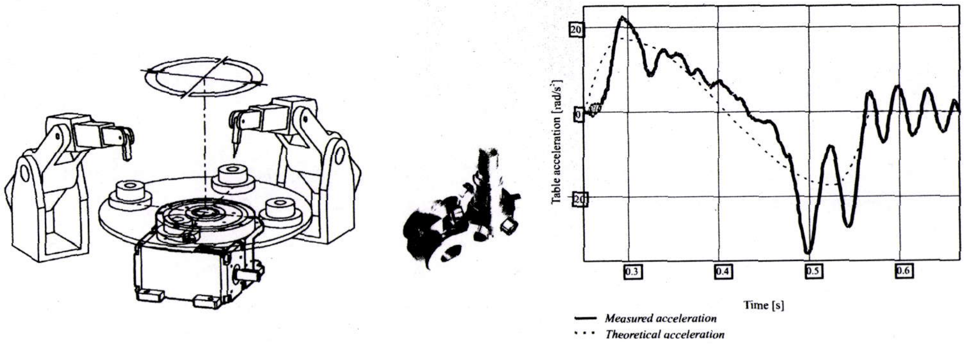
The rest of the paper is structured as follows:

- section 2 describes the mechanical system considered in the research;
- section 3 is devoted to a synthetic description of the feed-forward ANN;
- section 4 refers to the procedure of diagnosis for the simulated case: a description of the procedure, results of the analysis of parameters sensitivity and a discussion of classification results are given;
- section 5 concerns the experimental case: experimental device, tests performed and results achieved are presented in this order.
- section 6 resumes the main conclusions of the work.

2 The mechanical indexing system

A rotating table driven by an electrical motor through a gear speed reducer forms the mechanical system considered in the work [18]. The table is connected to the speed reducer through an indexing mechanism, formed by a spatial cam coupled with roller followers, to obtain intermittent motion (figure 1).

Such mechanisms are generally used by the manufacturing industry, in production lines which need an intermittent advance of the manufactured (figure 1). To amortize the costs of such systems, it's necessary to attain elevates levels of productivity. The maintenance purpose is to assure high productivity through the increment of the availability of the system and its conservation in the time.



a b c

Fig. 1: a)Example of industrial application of indexers. b)Spatial cam and roller follower of an indexing mechanism. c)Comparison between theoretical acceleration of the rotating table (continue line) and real acceleration (dotted line): vibrations of non negligible amplitude can be observed.

Frequently the motor used is a simple asynchronous triphase, with nominal speed of 1500 RPM. To reduce the rotation speed and to transfer motion to the indexing mechanism, it's necessary to introduce a transmission, which is often source of motion disturbance (distributed clearances and backlash often cause vibrations)[22, 14].

The indexing mechanism is the device that concurs to transform continuous rotatory motion in an intermittent rotatory one. The law of motion with dwells is obtained through the coupling between a cam (on the left in figure 2) and a series of rollers. The indexing mechanism is source of vibrations itself, caused by variations of the follower acceleration joined to the imprecisions in coupling between the cam and the follower.

Finally, also the load, commonly mounted directly on the indexing mechanism, can be source of differences between waited and obtained law of motion.

Experimental measures on systems with different sizes and series of indexers permitted to achieve a database of acceleration signals of the table. As example, figure 2 compares the real acceleration (continue line) versus the theoretical one (dotted line) in normal working condition. The amplitude of vibrations appears clearly non negligible.

3 Artificial feed-forward Neural Networks

Artificial Neural Networks (ANN) are structures that simulate the cognitive process that happens in mammals [5]. These systems of calculation differ from the traditional ones, based on the Von Neumann architecture, starting from the structure:

there is not, in the ANN, any division between elements dedicated to the calculation, to the memorization or to the allocation of the rules, but the tasks are distributed to all the neurons.

The base element of an ANN is the artificial neuron, which is connected to other equal neurons, in order to form an organic structure capable to carry out numerous tasks. The i-th artificial neuron receives n input $x_{i1}, x_{i2}, \ldots, x_{in}$. To every input it associates a weight ω_{ij}, that is a value that keeps account of the influence of that j-th input on the i-th neuron. The inputs, after weighted, are added. If the result of this sum exceeds a value θ, called activation threshold, the neuron proceeds with the calculation, applying to the result a function called "activation function" and obtaining the output of the neuron. This output is passed to all the neurons that follow.

$$y_i = \phi\left(\sum_{j=0}^{N} \omega_{ij} \cdot x_j\right) \tag{1}$$

The parameters that influence the operation of an artificial neuron are: the threshold, the weights and the activation function. The weights and the threshold are optimized during the training, whereas the activation function must be chosen during the definition of the net.

Usually, the activation function is chosen among the hard limit function (binary or bipolar), the linear function, the log-sigmoid function and the tan-sigmoid function. The choice of the activation function is made in relation to the type of code selected for the data (binary [0 1], or bipolar [-1 1]).

The training consists in the iterative modification of the weights, until the net is able to associate correctly input and target (output expected for that particular input). The training used for the implemented fault diagnosis tool is called "with supervisor", so to assign to every input the target, and then to estimate the classification error coming from the net. With proper mathematical operations, deriving from the rule of error back propagation, we calculate the amount of the correction to apply to every weight to minimize the error. The modification of the weights should terminate when the error coming from the net is cancelled. This is not always possible, both because it's difficult to find the combination of weights that cancels the error and because of the risk in the overfit of training data, due to the excessive quantity of training epochs, which may limit the generalization abilities of the net.

Several routes have been proposed in order to reach the best combination of weights. All of them start from the randomization of the initial values, then every one moves with different rules on the error surface.

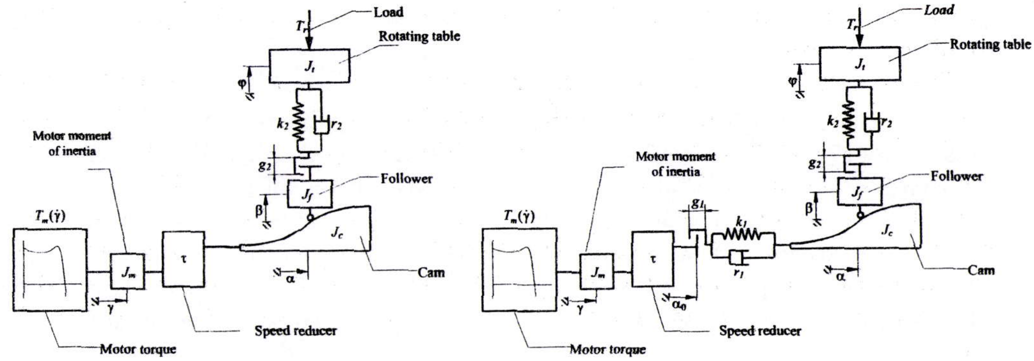


Fig. 2: a)Model A of the mechanism; b)Model B of the mechanism

4 Simulated case

4.1 The procedure

Elasto-dynamic models developed in a relevant research well simulate the real acceleration of the table, by taking into account clearances and elasticity in the system. Two different models are here considered: a one-degree of freedom model (afterwards called A) and a two-degrees of freedom model (B). Model A introduces an elasto-dynamic joint on the outside shaft of the indexer, while model B considers also an elasto-dynamic joint on the input shaft of the indexer. The comparison between measured and simulated acceleration for a great number of practical cases allows to say that model B better explains the most evident dynamic phenomena of systems with high operating speed and low rigidity, while model A is to be preferred when high loads and high inertias are involved.

The first phase of the developed procedure consists in the generation of the set of training examples by simulation, adopting model A and model B. Table 1 contains the nomenclature adopted throughout the work. The motion equation of the one-degrees of freedom model A is eq. (2), while motion equations of the two-degrees of freedom model B are (3).

$$\ddot{\varphi} = -\frac{1}{J_t}\left[T_L + k_2\left(\varphi - \beta\right) + r_2\left(\dot{\varphi} - \beta'\dot{\alpha}\right)\right] \tag{2}$$

$$\begin{cases} \ddot{\alpha} = & -\frac{1}{J_C + J_F\beta'^2}\left\{k_1\left(\tau\gamma - \alpha\right) + k_1\left(\tau\dot{\gamma} - \dot{\alpha}\right) + \beta'\left[k_2\left(\varphi - \beta\right) + \right.\right. \\ & \left.\left. + \, r_2\left(\dot{\varphi} - \beta'\dot{\alpha}\right) - J_F\beta'\beta''\dot{\alpha}^2\right]\right\} \\ \ddot{\varphi} = & -\frac{1}{J_t}\left[T_L + k_2\left(\varphi - \beta\right) + r_2\left(\dot{\varphi} - \beta'\dot{\alpha}\right)\right] \end{cases} \tag{3}$$

The acceleration of the rotating table ($\ddot{\varphi}$) is the parameter achieved to monitor the system. A Runge-Khutta of the fourth order method is used to solve the motion

γ, α	Motor and cam shaft rotation
β, φ	Follower and table rotation
J_m,J_C,J_F,J_t	Motor, cam, follower, platform moment of inertia
k_1,r_1,g_1	Stiffness, damping, backlash of joint 1
k_2,r_2,g_2	Stiffness, damping, backlash of joint 2
τ	Speed reduction ratio
T_L	Load torque

Table 1: Nomenclature adopted for the mathematical models.

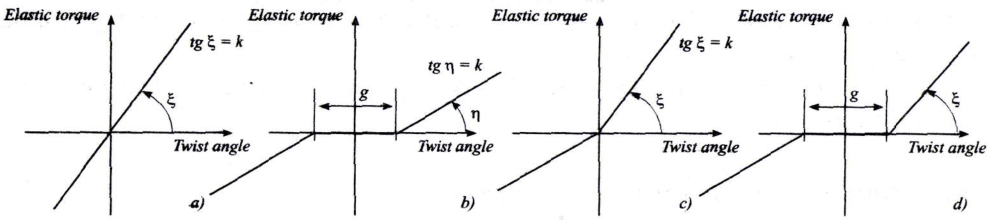


Fig. 3: Type of non-linearity and backlash considered . a) linear; b) linear with backlash; c) bilinear; d) bilinear with backlash

equations of the system, supposing constant the motor shaft speed. The values of the parameters of the model (k_1, r_1, g_1, k_2, r_2 and g_2) were previously established through an optimization procedure by comparing simulations with experimental measures of the acceleration $\ddot{\varphi}$. The non-linearities simulated are located at the joints and practically are implemented with non-linear characteristics (as shown in figure 3). In details, at joint 1 just the linear case is considered with or without backlash, at joint 2 the characteristic can be linear or bilinear, with or without clearance. The training examples were generated by solving the motion differential equations until transient condition is finished, several times with different values of the parameters of the elasto-dynamic model considered, in detail randomly changing the parameters in a range of $\pm 10\%$. The tangential acceleration on the table ($\ddot{\varphi}$) has been chosen as condition to monitor.

Step 2 of the procedure consists in the training of the NN. The NN used in this research is a multi-layer feed-forward NN, trained using the back-propagation algorithm (with momentum) and by adopting the early stopping method ([5]). The NN is trained using a first group of examples (the calibration set). The training is iteratively performed while monitoring the quality of classification on a second group (control set). The training is stopped when the error in the control set does not decrease any more. As a further control, at the end of the calibration, a prediction error is evaluated also on a third example set (the validation set). A three layer NN (with just one hidden layer) was chosen. It was assumed that a single hidden layer was sufficient; this is supported by the theorems of Cybenko and Funahashi which show that a single hidden layer is sufficient for the approximation of a polynomial function to arbitrary accuracy.

The performance rating of a solution is evaluated as the mean square errors on the control set, after the net training phase, so the performance improves if its value decreases.

In this context, parameters sensitivity analysis has been executed, to determinate the most suitable ANN for the diagnosis of the malfunctioning at issue, in a system for intermittent motion.

In details, parameters analyzed were different preproccessing techniques to create ANN inputs, random initialization, ANN hidden layer transfer function and number of nodes and the training rule.

In the course of the research, ANN were implemented with the Matlab Neural Network toolbox.

4.2 Analysis of parameters sensitivity

Comparison between FFT and PSD Different choices could be made for the pre-processing of the signal to supply as input to the ANN. Here the FFT (Fast Fourier Transform) and the PSD (Power Spectral Density) of the acceleration signal are compared. Experiments show that both FFT and PSD have an effective link with the machine operating condition; in fact, different cases have different transforms, and analogous cases, have similar transforms. Different classification and generalization tests were executed with FFT, filtering the signal with different low pass filters, with cut frequencies 100, 50, 25, 12.5, 6.3, 3.1 Hz, the aim being to identify the optimal one. Band pass width reduction involves both information lose on phenomena and dimensional reduction of the problem. The dimensional reduction is proportionally correlated with the input neurons number, that is the required number of FFT frequencies (fixed the test time at 8.2 seconds, the spectral resolution is 0.122 Hz and cannot be changed). For example, with a band with of 100 Hz, 800 values are required (800 neurons in input to the ANN), so the combination of the weights is searched in a 800 dimensional space. Analogous different types of PSD have been tried, changing the spectral resolution (approximately 0.5, 1, 2, 4, 8, 16, 32, 64 Hz), to see how much the PSD can be coarse, without loosing net abilities in classification and generalization.

For every case, 100 tests have been executed to achieve as much as possible insensibility to the initial randomization of the weights and both medium and best value obtained in every test has been reported on graph in fig. 4. Dotted lines visualize the mean values on 100 tests, while solid lines are related to the best performance both for the PSD and the FFT techniques.

One can notice that the nets which elaborate data obtained from the PSD achieve better performances than those trained with data preprocessed with FFT. When the PSD becomes excessive coarse, or when information on a sufficient number of

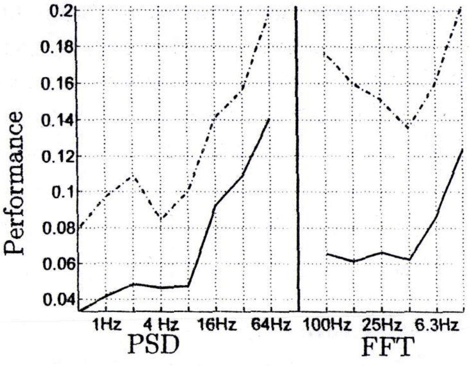


Fig. 4: Comparison between different types of PSD and FFT, changing spectral resolution and cut frequency. Mean (dotted) and best (solid) performance value on 100 tests.

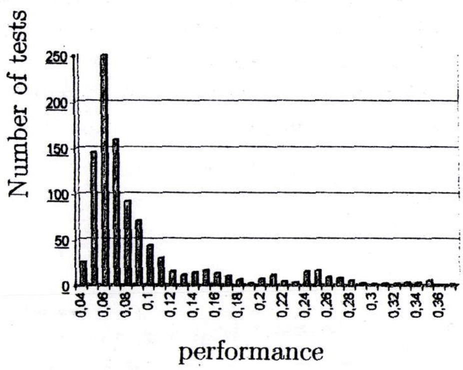


Fig. 5: Influence of net initialization on the final performance. Number of tests, among 1000, capable to approach the same performance.

frequencies is not supplied from FFT to the net, the performance get worse. In particular, lower limits to achieve acceptable performance are 12.5 Hz for the FFT and 8 Hz for the PSD. The PSD is clearly better than FFT not only for a better performance but also for the great reduction of the problem size; in fact, the FFT with band with 12.5 Hz requires 100 input nodes while the PSD with 8 Hz, as band resolution, requires only 12 nodes.

Initialization influence The weights initialization influences remarkably the classification and generalization performance of the ANN. Test were performed repeating the training 1000 times, starting from random values of the weight, to verify the influence of this parameter on results. Figure 5 shows the distribution of the number of tests versus the achieved performance (in abscissas). The histogram shows that the influence of the initialization can be remarkable. The more frequent value is about 0.06 (a low value for the considered case) and more than the 80% of the cases fall in an acceptable range around this value, so if a reasonable number of tests is executed, the probability to find a satisfactory result is high.

Number of hidden neurons An element that has high influence on the abilities of classification and generalization of the net is its architecture; an important parameter is the number of neurons of the hidden layer. In order to verify the hidden neuron number influence, several tests have been executed, searching for a possible optimal value. A configuration clearly better of the others has not been found (see fig. 6, and moreover, in this case, there are not great differences between the performances of the several net with a number of nodes equal or higher than 15.

Code influence In this section, the difference between the performance obtainable with bipolar or binary codifies is analyzed. Although often bipolar coding gives better results, thanks to the possibility to attribute no value to the zero, in our case the binary one allows the net to achieve better performances. This, probably, is due to the particular type of data used as input of the net, that are all positive values.

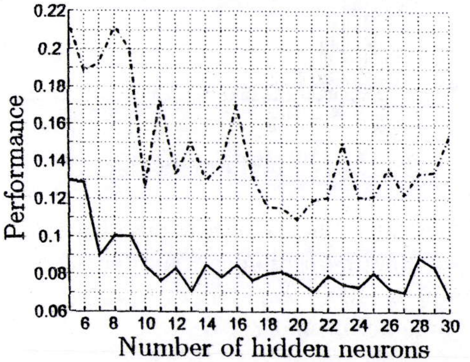


Fig. 6: Number of hidden neuron influence on net performance. The mean (dotted) and best (solid) performance value on 100 tests are plotted.

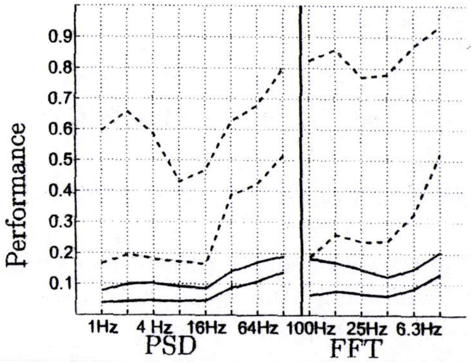


Fig. 7: Influence of binary and bipolar target representation on the performance of the net. Binary: solid line (mean and best), bipolar: dashed line.

Transfert functions Two configurations of the ANN in terms of hidden and output layer transfer functions were compared: the first one with logsigmoid functions for both the hidden layer and the output layer, the second with a logsigmoid for the hidden layer and linear for the output level. In figure 8 the achieved results are plotted. Better results are achieved with equal logsigmoidal functions.

Training rule The influence of the training rules on the net performances is at least tested. Several types of "trainers" have been compared. They are: variable learning rate backpropagation (traingdx and traingda), resilient backpropagation (trainrp), conjugate gradient (Fletcher Reeves updates - traincfg, Polak Ribire update - traincgp, Powell Beale restarts - traincgb, scaled conjugate gradient - trainscg), quasi Newton algorithms (BFGS algorithm - trainbfg, one step secant - trainoss), and Levenberg Marquard (trainlm). Although there are remarkable differences in the training durations, important differences in performances between the several tests have not be found (just trainrp algorithm has a very poor outcome. The best choice in terms of both performance and time cost was considered the traingdx algorithm.

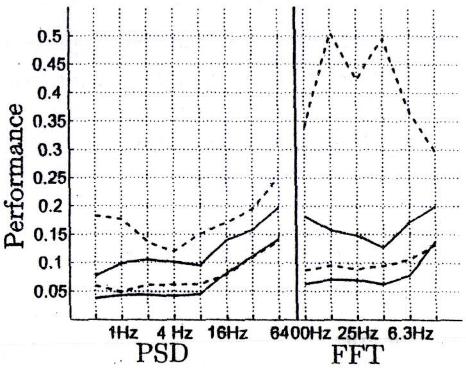


Fig. 8: Comparison, in terms of average and best performance, between equal activation functions (logsigmoid) for hidden and output layer (solid lines), and different activation functions (dashed lines).

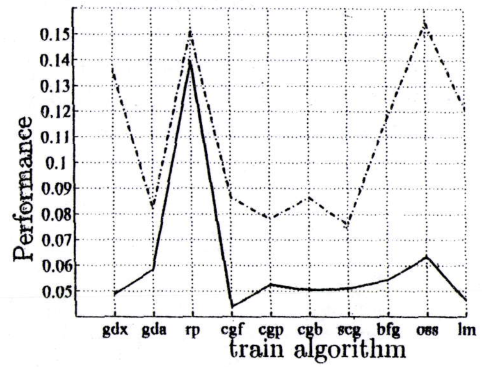


Fig. 9: Influence of the training algorithm on the net performance. In abscissas abbreviate algorithm names are referred. Each name must be preceded by "train".

4.3 Tests and numerical results

Six possible classes have been considered and five output nodes have been chosen, as described in Table 2.

Class	Model	Joint 1	Joint 2	Code
1	model A	-	linear without backlash	01111
2	model A	-	linear with backlash	01110
3	model B	linear without backlash	linear without backlash	11111
4	model B	linear without backlash	linear with backlash	11110
5	model B	linear without backlash	bilinear without backlash	11101
6	model B	linear without backlash	bilinear with backlash	11100

Table 2: Classes considered and related codification.

The main properties chosen for the network are: PSD of the simulated acceleration with 1 Hz of spectral resolution as ANN input, 20 hidden neurones, back-propagation with momentum and variable learning rate as training algorithm (traingdx), logsigmoidal activation functions and binary representation of the target. The examples generated by simulations are 126, equally subdivided for the training, for the control and for the validation. After 170 epochs the percentage error of the classification on all the three subsets considered is 0.3%. An experimental acceleration signal passed to the model has been classified in a very plausible way. The output of the network is: 0.9901, 0.9805, 1, 0.9989, 0.0032, 0.0011 [1 1 1 0 0], that is the network sentences that the system is simulated by the two degrees of freedom model, where joint 1 is linear without backlash while joint 2 is bilinear with backlash. This classification is very probable because, in the considered case, the backlash in the speed reducer is high and a roller follower of the indexer is cracked. Figure 10 compares the experimental acceleration with that simulated by model B with the configuration identified.

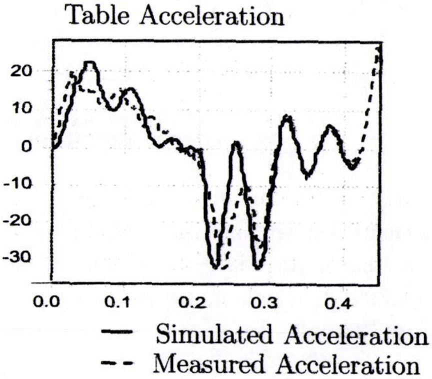


Fig. 10: Experimental acceleration versus that simulated and identified by the NN.

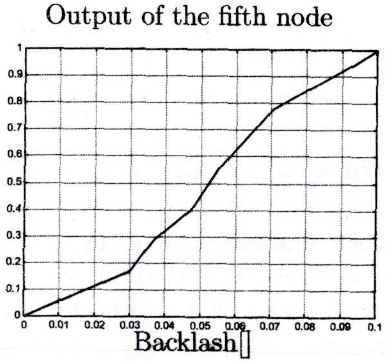


Fig. 11: Dependence between the NN output and the backlash amplitude. The fifth node value changes quite linearly with the backlash value in joint 2.

An interesting result is that the amplitude of the NN output is an index of the backlash in the joint. The output of the node related to the presence (1) or absence (0) of backlash in a joint changes quite linearly with the backlash amplitude (as shown in figure 11 for the backlash in joint 2 and for the fifth node). Such a property appears very useful for monitoring how clearances change during the working of the system, when wear phenomena occur.

5 Experimental case

5.1 The procedure

The experimental device to collect training data was formed by an indexing system (similar to the one described in section 3), an extensometric accelerometer with the respective instrumentation and a PC for data collection.

Three types of damaging of the indexing system have been tested: some rollers of indexer mechanism breaking, base screw loosening and backlash. In particular the backlash in the speed reducer has been made varying between 0 and 20 degrees, with a step of 5 degrees. The variable clearance was achieved by mounting a speed-reducer with very low nominal backlash and with a special joint (on purpose built) with adjustable backlash.

Each possible combination was considered, obtaining 20 different cases and for everyone, 10 tests for each case have been executed for statistical validity.

The code used to identify each case is formed by 3 bits: information about the indexer are contained in the first bit (0 intact, 1 damaged), about the screw in the second bit (1 lose, 0 tighted) and about the backlash in the third bit (0: zero backlash, 1: backlash 20/, and intermediate values proportional with backlash amplitude). For example code 1 1 0 means: index mechanism broken, screws loose and no backlash while code 1 1 0.5 means: index mechanism broken, screws loose and 10/ of backlash.

Similarly with the simulated case, the tangential acceleration on the table has been chosen as condition to monitor. The acceleration values have been acquired for 8,2 seconds, equal to one turn time of the rotating table; 1000 Hz has been chosen as sample frequency.

5.2 Tests and numerical results

In the light of the above analysis of parameters sensitivity, a net with 20 hidden neurons, logsigmoid for both transfer functions, traingdx as training rule, binary representation of the target and PSD with about 1 Hz of spectral resolution was chosen to create an automatic diagnostic system.

Firstly, the net ability of classification for remarkable cases was tested. Among the 10 collected series of data for each case, 7 have been used for training, and 3 for validation (to verify the net's ability in recognizing cases of the same type of the ones used in training). The net has demonstrated good ability in classification (the 100% of the cases) and any user, during test, would be able to assert, with certainty, in which case the system is.

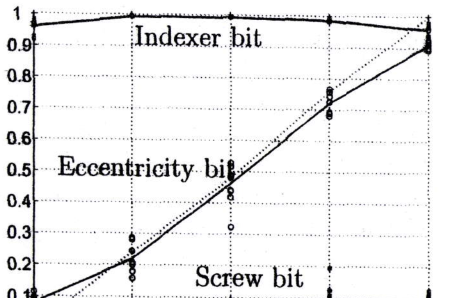


Fig. 12: Classification cases: 100, 101, 1 0 0.24, 1 0 0.48 and 1 0 0.76. Generalization cases: none. Expected results dot, achieved results dash or solid.

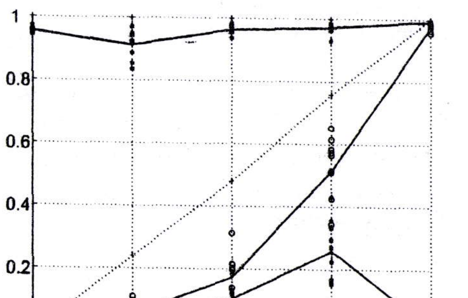


Fig. 13: Case *a*: Classification cases: 100, 101. Generalization cases: 1 0 0.24, 1 0 0.48 and 1 0 0.76. Expected results dot, achieved results dash or solid.

Figure 12 shows (for example) how the net classifies cases with indexer broken (value 1 for the first bit), tight screw (value 0 for the second bit) and variable backlash between zero and maximum value (value between 0 and 1 for the third bit). Circles in the graph represent the net result for each output node and adjoining circles correspond to repetitions of analogous cases. The mean value is considered and solid lines are plotted with the reasonable assumption of a linear transition between neighboring cases.

Afterwards, the generalization abilities have been tested, supplying a rising number of cases to the net in the training phase. Three tests were executed: by supplying only extreme cases (case *a*, fig. 13), an additional intermediate one (case *b*, fig. 14) and two intermediate cases (case *c*, fig.15).

Figure are relevant to the same damaging conditions considered in figure 12.

In case *a* it is evident that the error on generalization data for the third output node is high, in spite of this we have the information that the backlash is rising.

Cases *b* and *c* highlight that the training information density increment allows the ANN to become more and more expert both in classification and in generalization.

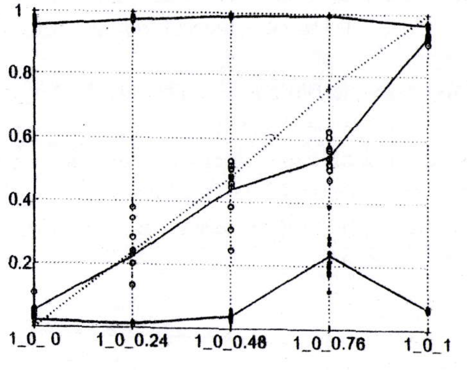


Fig. 14: Case *b*: Classification cases: 100, 101 and the middle one 1 0 0.48. Generalization cases: 1 0 0.24 and 1 0 0.76. Expected results dot, achieved results dash or solid.

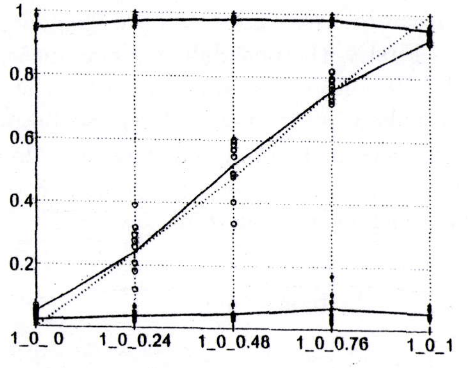


Fig. 15: Case *c*: Classification cases: 100, 101,1 0 0.24 and 1 0 0.76. Generalization cases: 1 0 0.48. Expected results dot, achieved results dash or solid.

6 Conclusions

The proposed ANN based approach to make a dynamical diagnosis on a mechanical system achieved good results, both for the simulated case and the experimental one. It is capable to classify various types of faults and has demonstrated generalization abilities, which improve if the density of the information given to the net is incremented. It is therefore clear that, in an industrial system case, with the passing of time and the increasing of the cases met, ANN progressively become more and more expert, both in the classification and in generalization.

When experiments with different causes of damage could be developed clearly the best solution is to train the ANN with experimental data, but when these test are not possible to use training data simulated by the use of mathematical models could be an effective alternative.

References

[1] R. Bussola, R. Faglia, G. Incerti, P.L. Magnani, M. Tiboni (1999). Indexing Cam Mechanisms: a Complex Mathematical Model for Simulation. In *Proceedings of the X World Congress on the Theory of Machines and Mechanisms (IFToMM)*, Oulu, June 1999.

[2] Zhi-Qiang Hou, A. Lucifredi (1995). On the use of Neural Networks for identification of linear and nonlinear systems. In *Meccanica 30, Klower Academic Publisher*, pp. 377-388, 1995.

[3] K. Worden, G.R. Tomlinson (1992). Classifying linear and nonlinear systems using neural networks. In *Proceedings of the 17th International seminar on Modal Analysis, Leuven*, pp. 903-922, 1992.

[4] R. Bussola, M. Tiboni (1999). Parametrical identification for a complex mathematical model of indexing cam mechanisms by means of a Genetic Algorithm. In *Proceedings of the EUROGEN99: Short Course on Evolutionary Algorithms in Enginering and Computer Science*, Jyväskylä, June 1999.

[5] Haykin S.(1999). Neural Network: A Comprehensive Foundation, Prentice-Hall, New Jersey, 1999.

[6] A.V.Oppenheim and R.W.Schafer (1989). Discrete-time signal proccesing, Prentice Hall,London,1989.

[7] B.A.Paya, I.I.Esat and M.N.M.Badi (1997). Artificial Neural Networks Based Fault Diagnostic of Rotaring Machinery using Wavelet Transforms as a Preprocessor, title = "Mechanical Systems and Signal Processing", vol 11, Accademic Press Limited,1997.

[8] A.T.Vemuri, M.M.Polycarpou and S.A.Diakourtis (1998).IEEE Transaction on Robotics and Automation,Neural Networwk Based Fault Detection in Robotic Manipulators,Publisher Item Identifier 1042-296X/98,pp. 342-348,1998.

[9] J.N.K.Liu and K.Y.Sin (1997). IEEE Transaction on Neural Networks,Fuzzy Neural Networwk for Machine Maintenance in Mass Transit Railway System, Publisher Item Identifier 1045-9227/97,pp. 932-941,1997.

[10] D.J.H.Wilson, G.W.Irwin and G.Lightbody (1999). IEEE Transaction on Neural Networks, RBF Principal Mainfolds for Process Monitoring,Publisher Item Identifier 1045-9227/99, pp. 1424-1434,1999.

[11] P.J.Edwards, A.F.Murray, G.Papadopoulus, A.R.Wallace, J.Barnard and G.Smith (1999). IEEE Transaction on Neural Networks,The Application of Neural Networks to the Papermaking Industry,Publisher Item Identifier 1045-9227/99,pp. 1456-1464,1999.

[12] S.D.G.Smith, R.Escobedo, M.Anderson and T.P.Caudell (1997). IEEE Transaction on Neural Networks, A Deployed Engineering Design Retrival System Using Neural Networks, Publisher Item Identifier 1045-9227/97,pp. 847-851, 1997.

[13] V.R.deAngulo and C.Torras (1997).IEEE Transaction on Neural Networks, Self Calibration of a Space Robot, Publisher Item Identifier 1045-9227/97, pp. 951-963,1997.

[14] R. Bussola, R. Faglia, P.L. Magnani, M. Tiboni,(1999). Proceedings of the X World Congress on the Theory of Machines and Mechanism (IFToMM), Indexing Mechanism: a Complex Mathematical for Simulation, 1999.

[15] Zhi-Qiang Hou, A. Lucifredi (1995). Meccanica 30, On the use of Neural Networks for identification of linear and nonlinear systems, pp. (377-388), "Klower Academic Publisher",1995.

[16] K. Worden, G.R. Tomlinson (1992).Proceedings of the 17th International seminar on Modal Analysis, Classfying linear and nonlinear system using neural networks, pp. 903-922.

[17] R. Bussola, M. Tiboni (1999). Proceedings of the EURO-GEN99: Short Course on Evolutionary Algoritms in Engineering and Computer Science, Parametrical identification for a complex mathematical model of indexing cam mechanism by means of a Genetic Algorithm,1999.

[18] P.L.Magnani and G.Ruggieri (1986). Meccanismi per Macchine Automatiche,UTET,Torino,1986.

[19] E.C. Fitch (1992). Proactive Maintenance for Mechanical Systems, Elsevier Science Publisher, Oxford UK,1992.

[20] S.S. Rao,(1995). Machanical Vibration,Addison-Wesley Publishing,1995.

[21] A.V. Oppenheim and R.W. Schafer (1975). Digital signal processing, Prentice Hall,London,1975.

[22] J. Reeve (1995). Cams for Industry, Mechanical Engineering Publications Limited, London,1995.