Computer Assisted Mechanics and Engineering Sciences, 14: 557–567, 2007.

Rotating machinery diagnostics based on NARX models

Jarosław Bednarz, Tomasz Barszcz, Tadeusz Uhl
AGH University of Science and Technology, Department of Robotics and Mechatronics
al. Mickiewicza 30, 30-059 Kraków, Poland

(Received in the final form September 18, 2007)

Rotating machines are often described using linear methods with acceptable accuracy. Some malfunctions, however, are of non-linear nature. Accurate detection and identification of such malfunctions requires more accurate methods. One of such methods can be NARX — Non-linear AutoRegressive model with eXogenous input. The paper presents how NARX models can be applied for modeling rotating machinery malfunctions. Idea of the diagnostic algorithm based on such modeling is presented. Proposed algorithm was verified during research on a specialized test rig, which can generate vibration signals. The paper presents results of application of NARX models for detection of typical rotating machinery failures and the variations of NARX model parameters due to propagation of damage. In the paper authors present also a blade crack detection using the NARX models. The last chapter of the paper discusses the applicability of this method for damage detection in real machines.

Keywords: rotating machinery diagnostics, blade crack detection, neural networks, NARX models

1. INTRODUCTION

Rotating machines play a vital role in modern economics. Most industrial processes where energy is processed are based on rotating machinery. Thus, it is increasingly important to maintain those machines in the good technical state. Main drivers for final users are:

- avoidance of catastrophically failures
- decrease of maintenance costs
- increase of availability

This needs, in turn, create strong demand for diagnostic techniques. Theoretical works are performed since decades, starting from simplified, linear rotor models. With advances in rotordynamics research new processes were identified and described. Extensive review of theoretical description of rotordynamics phenomena can be found e.g. in [11]. In many cases those phenomena are of non-linear nature. Conclusion stating that a rotor and its support consisting of hydrodynamic bearings, formulate a highly nonlinear closed-loop system is also known [9]. Additionally, several malfunctions have inherently nonlinear nature. Good examples are mechanical looseness or rotor rub. Such problems were also investigated and solved in the industrial practice. Interesting survey of rotating machinery malfunctions can be found e.g. in [4].In the general case, to model a rotating object (with or without malfunctions), the system input may be stated as the forces relating to rotor imperfections such as unbalance or overloads connected to external forces acting upon the shaft [2]. The state vector includes velocities and displacements of nodes, where lumped physical parameters are focused. The equations strongly depend on parameters, especially rotational speed Ω.

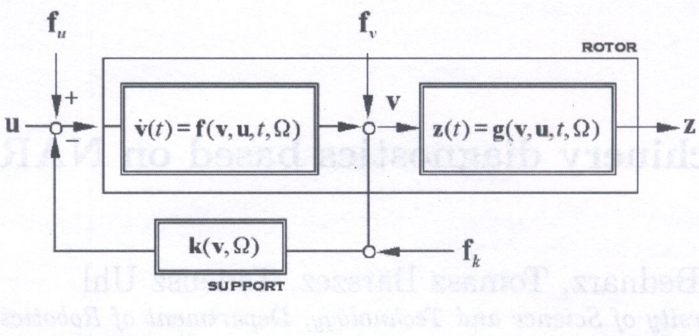

Fig. 1. A block scheme of rotor and support equations affected by faults in the form of state-space equations

Parameters of a diagnostic model to be identified from measured data should be very sensitive to the symptoms of malfunctions of rotating machinery. The problem of faults description (Fig. 1) of supported rotors can be formulated based on nonlinear state-space equations as follows,

$$\begin{aligned} \dot{v}(t) &= f\left(t, v(t), u(t), w(t), f_{k,u}(t), \theta\right), \\ z(t) &= g\left(t, v(t), u(t), w(t), f_v(t), \theta\right). \end{aligned} \tag{1}$$

The parameters of the vector θ are physically significant and their changes correspond to faults (malfunctions) to be detected and isolated. A major problem with rotordynamics modeling is that equations strongly depend on parameters, especially rotational speed. In practice, a large group of machinery is operating at a constant speed for very long periods. On the one hand, it simplifies the problem by reducing the number of degrees of freedom. On the other hand, one needs to deal with nonlinear equations with less input data available. The general problem presented above is often approached using linear system identification methods. In many cases such an approach yields good results, allowing detection and identification of machinery faults. In some cases, linearization of a nonlinear model can also bring useful results [1]. In other cases, as mentioned previously, nonlinearities are inherent and linear models can be only used in a limited scope. Mechanical looseness is the inherently nonlinear phenomenon. It is encountered when a stationary machine part (e.g. bearing pedestal) becomes loose. In such a case, the effective rotor stiffness is reduced. This often results in rotor resonance shift into a frequency which is an even multiple (or fixed fraction) of rotating speed. Additionally, due to mechanical looseness synchronous motion of a part may become truncated, e.g. due to hammering of parts. Truncated sine waves exhibit a series of running speed harmonics. Those harmonics, in turn may induce resonances of other parts of the machine. Examples of practical cases of looseness can be found in [4]. Proposal of analytical model of this malfunction was given in [11]. The model of the rotor lateral mode with additional terms, due to coupling with the stationary part can be presented as follows:

$$\begin{aligned} &M\ddot{W} + (D + \nu D_f)\dot{W} + KW + \nu K_f(W - c) + K_{XY}W\cos 2\psi - MW\dot{\psi}^2 \\ &\qquad = mr\Omega^2\cos(\Omega t + \delta - \psi) + P\cos(\gamma - \psi), \\ &MW\psi^2 + (D + \nu D_f)W^2\dot{\psi} - K_{XY}W^2\sin 2\psi - 2MW\dot{W}\dot{\psi} \\ &\qquad = W\left[mr\Omega^2\sin(\Omega t + \delta - \psi) + P\sin(\gamma - \psi)\right] + \nu(W + R_r)F_t, \end{aligned} \tag{2}$$

where

$$\nu = \begin{cases} 1 & \text{if } |W| \geq c, \\ 0 & \text{if } |W| < c. \end{cases} \tag{3}$$

The main assumption of the model is discontinuous (and thus nonlinear) change of the stiffness. The same model can be applied for looseness and for rubbing. In the first case, change of stiffness is caused by pedestal/surface contact. In the latter one – by rotor/stator contact. In some systems, when stiffness (and sometimes damping) varies between extreme discrete values, the system given in Eqs. (2)–(3) may become chaotic. Basics of chaotic motion were described in [10]. The suspended rotor with looseness or rub is one of examples, where such a chaotic motion can take place. As shown in [5], the condition for such behavior is the discontinuous nonlinearity of the system, and its high sensitivity to deterministic external excitation. To model such nonlinear objects, neural networks can be applied.

2. Fundamentals of NARX

The most general structure of nonlinear 'black-box' model is a neural input-output model. It is recommended for modeling cases without a priori knowledge of structure and its nonlinearities. The neural network is a set of parameters called weights and biases. Various architectures of neural networks are described [7]. The most often applied network structure is the multi-layer perceptron (MLP). Example of structure of two layers MPL is presented in Fig. 2.

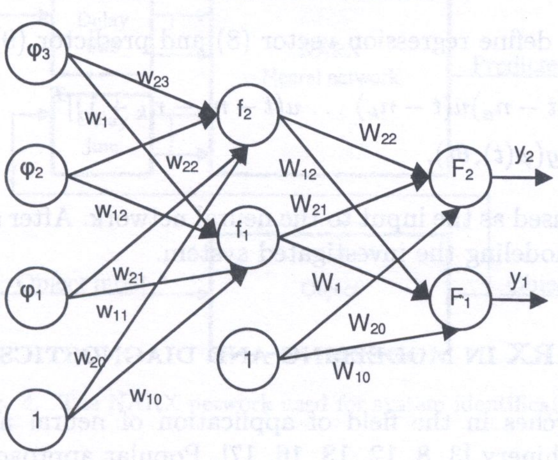

Fig. 2. The structure of 2 layer Multi Layer Perceptron

The general case of this architecture of the neural network can be given by the expression

$$\hat{y}(t) = g[\varphi, \theta] = F_i \left[\sum_{j=1}^{nH} W_{i,j} f_j \left(\sum_{l=1}^{n\Gamma} w_{j,l}\varphi_l + w_{j,0} \right) + W_{i,0} \right]. \tag{4}$$

The predictor $\hat{y}(t) = g[\varphi, \theta]$ consists of the past outputs and/or the past inputs and predicted output $\hat{y}(t)$ where θ denotes the parameter vector, which contains all the adjustable parameters of a network. Here, biases were presented as weights with second index 0. Usually sigmoid/tansig activation functions are applied in the hidden layer neurons, whereas linear – in the output layer. The structure given in Eq. (4) was considered. The weights, referred to as q, or w and W, are adjusted during the training process based on a training set of inputs and outputs. The learning criterion is the least mean square error between the given output and the predicted output. The formula for prediction error is given by equation

$$PE = \frac{1}{2N} \sum_{r=1}^{N} [y(t) - \hat{y}(t|\theta t)]^T [y(t) - \hat{y}(t|\theta t)]. \tag{5}$$

The weights are found according to the learning algorithm. The basic one is based on the back-propagation. Detailed description this algorithm and other ones are given in e.g. [6]. The multi-layer perceptron can be applied to identify or model a nonlinear dynamic system [7]. The structure, which will be investigated here, is referred to as NARX — Nonlinear Auto Regressive model with eXogenous input (6, 7). The NARX model [6] represents a wide class of non-linear systems, and many well-known non-linear input-output models are specific cases of this model.

$$y(t) = \sum_{m=1}^{M} y_m(t), \tag{6}$$

$$y_m(t) = \sum_{p=0}^{m} \sum_{k_1, k_{p+q}=1}^{K} c_{p,q}(k_1, \ldots, k_{p+q}) \times \prod_{i=1}^{p} y(t-k_i) \prod_{i=p+1}^{p+q} u(t-k_i), \tag{7}$$

where $y_m(t)$ is the mth-order output of system, $p+q=m$, $k_i = 1, \ldots, K$,

$$\sum_{k_1, k_{p+q}=1}^{K} (\cdot) = \sum_{k_1=1}^{K} (\cdot) \cdots \sum_{k_{p+q}=1}^{K} (\cdot).$$

For such a structure, we define regression vector (8) and predictor (9) as

$$\varphi(t) = [y(t-1) \ldots y(t-n_a)u(t-n_k) \ldots u(t-n_b-n_k+1)]^T, \tag{8}$$

$$\hat{y}(t|\theta t) = \hat{y}(t|t-1,\theta) = g(\varphi(t), \theta t). \tag{9}$$

The regression vector is used as the input to the neural network. After successful learning process, the network is capable of modeling the investigated system.

3. Application of NARX in modelling and diagnostics

There are numerous researches in the field of application of neural networks for modeling and diagnostics of rotating machinery [3, 8, 12, 13, 16, 17]. Popular approach is the application of the neural network in the process of classification. The idea of this approach is presented in Fig. 3.

In such a configuration, the output layer is composed of neurons having unipolar step activation functions. Activation of an output occurs when the network classifies the state as certain (correct or one of the failure types). There are many possibilities when inputs are concerned. The most successful approaches apply spectral lines or selected parameters of vibration signals. Usage of perceptrons in presented approach has some disadvantages:

- low immunity to noise,
- high number of inputs (especially in case of feeding spectral lines as inputs),
- inability to detect combined failures,
- learning set is necessary for good results.

The necessity of having the learning set containing all failures one wants to detect is especially problematic. In practice, behavior of the machine is extremely rich and depends on a multitude of factors. In consequence, collection of a reasonable learning set is rarely possible. On the other hand, NARX networks were applied for modeling and identification of dynamic systems [13], yielding good results. Typical architecture of such application is presented in Fig. 4.

Application of such a network for diagnostics of rotating machinery needs to solve two problems:

- unavailability of input signal
- interpretation of results

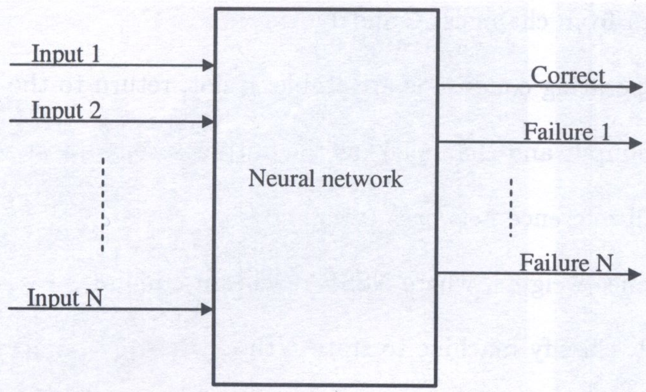

Fig. 3. The neural network used for classification

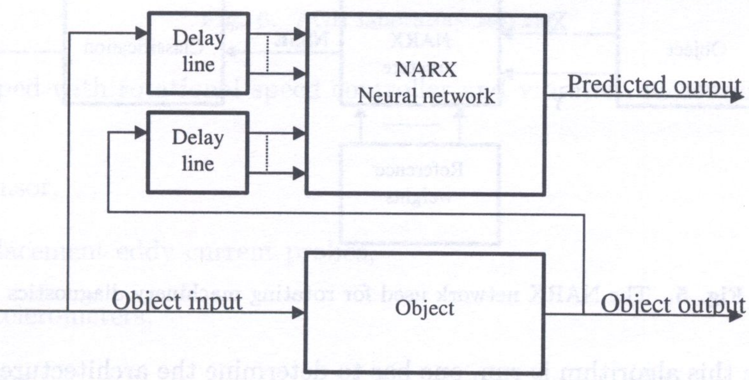

Fig. 4. The NARX network used for system identification.

The input signals in rotating machinery are the forces acting on the structure. The majority of this force, according to Eq. (2), is the rotor unbalance force. After being transmitted through the object it produces vibrations, which are observed as system outputs. The rotor unbalance force is immeasurable and – even worse – it is nonstationary, for it depends on rotational speed, rotor temperature etc... In such a case either no input signal will be used, resulting in degradation of quality of the model, or another signal can be used as the input. In this approach, we propose to use one of vibration signals as the input signal. In such an approach the NARX network will approximate a nonlinear filter, transforming vibration signal from one channel to the other one. Interpretation of results causes problems, because neural network parameters (i.e. weights) do not have any physical interpretation. In the presented approach the output from the network is the predicted vibration signal (or signals, in more general case, but we will deal with only one output at the moment), so there is no direct information about any malfunction. We propose to train the network and next calculate only "distance" between current set of input/ output data and the set used to train the network. In this case by "distance" authors mean the quality of neural network validation with new set of data. Good scalar estimate of this value is normalized sum of squares of prediction errors (NSSE – expressed by Eq. (10)), when new data is used to validate the NARX network trained with reference data.

$$PE = \frac{1}{2N}\sum_{t=1}^{N}[y(t) - \hat{y}(t|\theta t)]^T[y(t) - \hat{y}(t|\theta t)] + \frac{1}{2N}\theta^T D\theta. \tag{10}$$

Thus, the output will be the set of numbers, showing how much the current state of the machine differs from selected reference states known previously. The proposed approach is presented below:

- acquire vibration data from channels X and Y
- check, whether the operating conditions are stable; if not, return to the beginning
- use channel X as the input and channel Y as the output; prepare test set
- calculate NSSE for all reference networks (weights)
- find set Nth of reference weights, where NSSE reaches minimum
- if min(NSSE) < limit, classify machine to state Nth

The architecture of proposed algorithm is shown in Fig. 5.

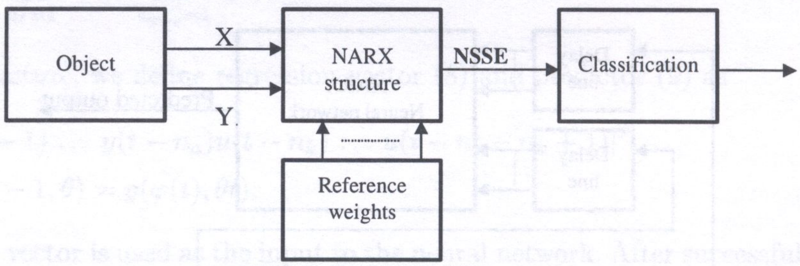

Fig. 5. The NARX network used for rotating machinery diagnostics

Note, that before this algorithm is run, one has to determine the architecture of the network, i.e. order of the NARX network (number of past inputs and outputs) and number and type of neurons in the hidden layer. Also, at least one state must be known to the network (one reference weights vector must be known). Presentation of results can be given in various formats. We propose tabular view with the NSSE to known reference sets. If the distance for an Nth is smaller than a predefined limit, the machine is assumed to work in this closest Nth state. If the distance is greater than the limit from all known reference vectors, the unknown state should be displayed, notifying the expert to investigate the machine and probably create a new reference set.

The algorithm was prepared for rotating machinery, operating for long periods with constant rotational speed. This is common case for e.g. power generation machinery, when lack of transient states makes diagnostics difficult. On the other hand, we ignore inputs, when the rotational speed is other than nominal. The algorithm was prototyped based on the data from a real test rig. Following chapters present this experiment.

4. Rotating shaft – experimental rig installed at AGH

The rotordynamics test rig (Fig. 6) was designed and installed in Department of Robotics and Mechatronics in University of Science and Technology. Its main goal is research of diagnostic techniques for rotating machinery. Additionally it is used for other research work (e.g. Operational Modal Analysis).

The rotor – bearing system is mounted on the heavy steel plate. The rotor is driven by the 1.2 kW AC motor, controlled by the converter. The converter controls the rotational speed set manually or from a PC through a serial link. Various transient states can be easily tested.

The driven system can consist of one 1200 mm or two 600 mm rotors, mounted on bearings. The rig has exchangeable bearings, rolling and sliding ones. One of bearing supports can introduce controlled misalignment. Up to four disks can be mounted on rotors, to introduce static or dynamic unbalance. It is possible to introduce looseness and rotor rub.

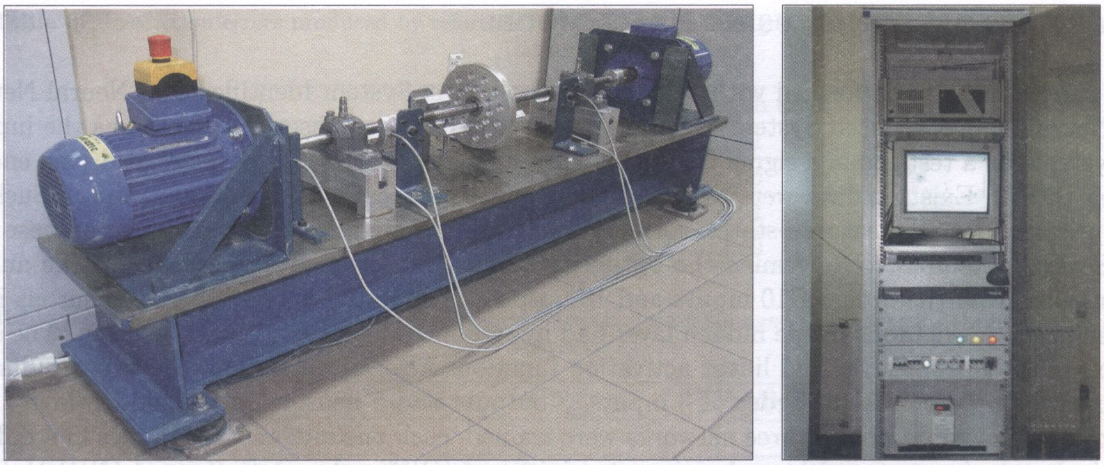

Fig. 6. AGH laboratory test rig

The rig is equipped with rotational speed controller and vibration measurement system, which consists of:

- phase marker sensor,
- 8 vibration displacement eddy-current probes,
- 2 three-axial accelerometers,
- signal conditioning,
- analog input measurement card,
- acquisition computer,
- software.

Vibration displacement sensors are mounted on pedestals, two sensors per one. This allows to measure shaft vibration in any position, not only at the bearings. Acceleration sensors are mounted on bearing pedestals.

The sampling frequency can be configured with the default value of 2 kHz. Apart of storage of raw vibration data, the measurement systems calculate following parameters of the vibration signals:

- root mean square,
- peak-peak amplitude,
- amplitude and phase of the fist harmonic,
- amplitude and phase of the second harmonic,
- DC value (for eddy-current probes).

Stored data can be exported for future processing (e.g. in Matlab environment). It can be also presented in one of following plots: time trend, waveform, spectrum, cascade, orbit, Bode, polar.

The measurement system is very flexible and can be easily adapted to the particular experiment. It is for example possible to increase the sampling frequency (up to 20 kHz per channel) or to add new signal parameters, like other harmonics, or sub synchronous component.

5. Failures detection using NARX model

Data processing was performed with Matlab and toolboxes: System Identification, Neural Network and Neural Network Based System Identification v. 2.0 [15]. Keyphasor was chosen as the input to the system. As a test signal a signal from eddy current sensor (vertical directions) at driven end was chosen (DEZ). Existing data were divided into separate sets, each having 2500 samples. Such sets were prepared for all measured states: correct, unbalance, misalignment and two cases of loose bearing. Several attempts to determine the optimum network structure were performed. The analysis was started from the order of 10 inputs and 10 outputs. The initial network had 10 neurons in the hidden layer. All neurons in the hidden layer had hyperbolic tangent activation function. The single neuron in the output layer had linear activation function. After optimization, the best results were obtained for network having order of 3 inputs, 3 outputs and 5 neurons in the hidden layer. After network optimization phase, three networks were trained, each one modeling dynamics in a different technical state (correct and loose bearing at driven end (DE) and non-driven end (NDE)). Those trained networks were later used as reference networks. The goal of the algorithm is to detect, if the set of currently acquired data can be classified to one of known states. To verify this idea, 3 sets of validation data (each one consists of 5000 input and 5000 output samples), each taken from measurement with different malfunction present, were presented to each network. The measure of distance between real data and predicted output was normalized sum of squared prediction error (NSSE). Thus, 9 estimates were obtained for every channel. Table 1 presents results for the channel DEZ.

The table shows clear difference between data from the same state, for which the network was trained and other data. However, the network trained from data with loose bearing at NDE shows worse ability to distinguish between the data. There is some difference between data from the correct state and data with malfunction, but both malfunctions are practically impossible to distinguish. After these tests we decided to use the reference network for object undamaged for detecting another malfunctions – unbalance and misalignment. The results of this analysis are shown in Table 2.

The next step of our researches was focused on checking the influence of damage propagation for NSEE value. We made this experiment for unbalance and misalignment. We introduce three degrees of these failures: small, medium and large. We trained a new neural network for analysis of this experiment results. The results of these experiments are presented in Tables 3 and 4 for accelerometers and in Tables 5 and 6 for eddy-current sensor.

Experiments results clearly show that NARX models can be successfully implemented in rotating machinery diagnostics. Value of prediction error for NARX approach increases for increasing damage. This mean that we can calculate estimated time for normal operating of machine. Presented results proved applicability and advantages of NARX models in model based structural health monitoring of rotating machinery. Next works will be consider on build a algorithm which allows to build a most fitted neural network for every kind of machines.

Table 1. Prediction errors produced by three reference networks for three validation data sets

Reference network	Data OK	Data from loose DE	Data from loose NDE
No malfunction	1.52	3.52	1.77
Loose bearing DE	5.12	2.29	2.87
Loose bearing NDE	1.75	3.05	1.50

Table 2. Prediction errors produced by reference network (for undamaged structure) for three validation data sets

Data type	Prediction error
No malfunction	1.52
Data from misalignment	7.86
Data from unbalance	9.20

Table 3. Prediction errors produced by reference network (for undamaged structure) for three validation data sets

Data type	Prediction error
No malfunction	1.41
Data from small unbalance	1.66
Data from medium unbalance	1.91
Data from large unbalance	2.49

Table 4. Prediction errors produced by reference network (for undamaged structure) for three validation data sets

Data type	Prediction error
No malfunction	1.41
Data from small loose at NDE bearing	1.49
Data from medium loose at NDE bearing	1.70
Data from large loose at NDE bearing	1.75

Table 5. Prediction errors produced by reference network (for undamaged structure) for three validation data sets

Data type	Prediction error
No malfunction	2.09
Data from small unbalance	2.37
Data from medium unbalance	2.64
Data from large unbalance	3.03

Table 6. Prediction errors produced by reference network (for undamaged structure) for three validation data sets

Data type	Prediction error
No malfunction	2.09
Data from small misalignment	5.05
Data from medium misalignment	5.43
Data from large misalignment	8.71

6. Blade crack detection using NARX model

Recently Application of the NARX approach to blade crack detection was investigated. The test rig was modified by mounting the blades at the end of the shaft. Then a crack of a blade was introduced (Fig. 7).

In the NARX approach data processing was performed in Matlab and dedicated toolboxes: System Identification, Neural Network and Neural Network Based System Identification v. 2.0 [15]. As the input to the system the Keyphasor was chosen. Two signals: a signal from eddy current sensor (vertical directions) at the driven end and a signal from accelerometer (vertical directions) at the driven end were chosen as test signals. Existing data was divided into separate sets, each having 10000 samples. Such sets were prepared for all the measured states: correct as well as with a cracked blade, separately for accelerometers and eddy-current sensors. Several attempts to determine the optimum network structure were performed. The analysis started from order of 10 inputs and 10 outputs. The initial network had 10 neurons in the hidden layer. All neurons in the hidden layer had hyperbolic tangent activation function. The single neuron in the output layer had linear activation function. After optimization, the best results were obtained for network having order of 3 inputs, 3 outputs and 5 neurons in the hidden layer. After network optimization phase, four

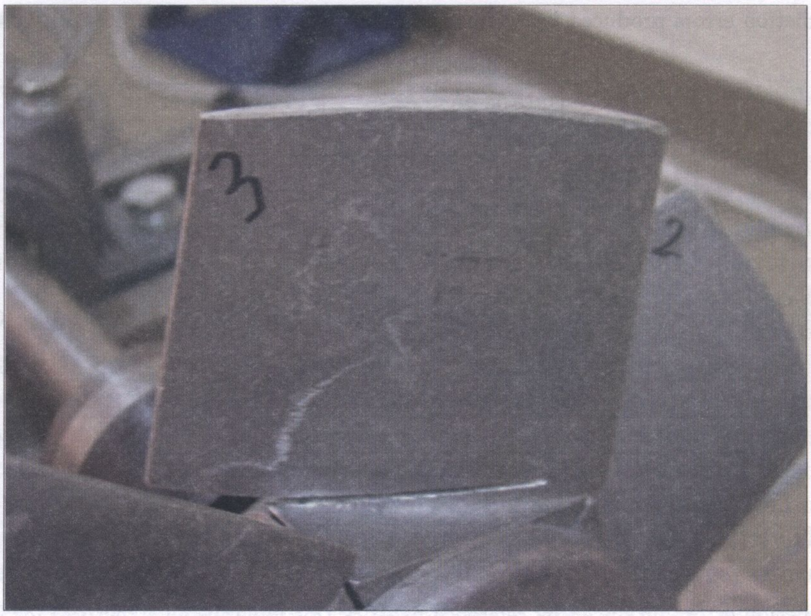

Fig. 7. Simulation of crack at blade

networks were trained (two for accelerometers and two for eddy-current sensors) — each one modeling dynamics in a different technical state. Those trained networks were later used as reference networks. The goal of the investigation was to detect whether the set of currently acquired data could be classified to one of the known states. To verify this idea, 4 sets of validation data (each set consisted of 35000 input and 35000 output samples), resulting from measurements in the presence of different malfunctions, were used as inputs to the considered networks. As the measure of instance between real data and predicted output the normalized sum of squared prediction error (NSSE) was assumed. Thus, 4 estimates were obtained for every channel. In Table 7 there are presented results for the accelerometer while in Table 8 for the eddy-current sensor.

Table 7. NARX results– Accelerometers (NSSE)

Neural network \ Data	Undamaged	Blade crack
Undamaged	11.00	12.60
Blade crack	12.42	11.00

Table 8. NARX results – Eddy-Current Sensors (NSSE)

Neural network \ Data	Undamaged	Blade crack
Undamaged	1.9555	2.0337
Blade crack	85.0570	4.6917

7. Conclusion

The paper presents application of NARX neural network to diagnostics of rotating machinery. The proposed method was verified at the test rig, where correct state and two cases of mechanical looseness failures were introduced. Performed research showed ability of proposed algorithm to detect introduced malfunctions. The algorithm was designed to work in the steady state only, for it should be applicable for machines working on a constant rotational speed for long periods of time. Proposed algorithm is very processing power demanding, but does not exceed capabilities of modern computers.

ACKNOWLEDGMENTS

Scientific research was financed from Polish means for science as the research project 4T07B 00430: 'Application of OMAX method to power plant diagnostics' (2006–2008).

REFERENCES

[1] T. Barszcz. Nonlinear system identification for diagnostic of turbine control system. In: *Proc. Of 11th IEEE International Conference Methods and Models in Automation and Robotics*, Międzyzdroje, Poland, Aug. 29–Sept. 1, 2005

[2] T. Barszcz, P. Czop, T. Uhl. System identification and its limitations relating to the diagnosis of rotating machinery faults. In: *Proc. of 10th IEEE International Conference Methods and Models in Automation and Robotics*, Miedzyzdroje, Poland, Aug. 30–Sept. 2, 2004.

[3] Ch. Chen, Ch. Mo. A method for intelligent fault diagnosis of rotating machinery. *Digital Signal Processing*, **14**: 203–217, 2004.

[4] R.C. Eisenmann. *Machinery Malfunction Diagnosis and Correction.* Hewlett Packard Professional Books, 1997.

[5] P. Goldman, A. Muszynska. Chaotic behavior of rotor/stator systems with rubs. *Journal of Engineering for Gas Turbines and Power*, **116**: 692–701, 1994.

[6] X.J. Jing, Z.Q. Lang, S.A. Billings. New bound characteristics of NARX model in the frequency domain. *International Journal of Control*, **80**: 140–149, 2007.

[7] J. Korbicz, A. Obuchowicz, D. Uciński. *Artificial Neural Networks* (in Polish). Akademicka Oficyna Wydawnicza, Warsaw, 1994.

[8] J.M. Kościelny. *Diagnostics of Automatic Industrial Processes* (in Polish). EXIT, Warsaw, 2001.

[9] E. Kraemer. *Dynamics of Rotors and Foundations.* Springer-Verlag, 1993.

[10] E.N. Lorenz. Deterministic non periodic flow. *Journal of Atmospheric Sciences*, **20**, 1963.

[11] A. Muszynska. *Rotordynamics.* Taylor and Francis, Boca Raton, 2005.

[12] N.S. Vyas, D. Satishkumar. Artificial neural network design for fault identification in a rotor-bearing system. *Mechanism and Machine Theory*, **36**: 157–175,2001.

[13] K.S. Narendra, K. Parthasarathy. Identification and control of dynamical systems using neural networks. *IEEE Trans. Neural Networks*, **1**(1): 4–27, 1990.

[14] M. Norgaard, O. Ravn, N.K. Poulsen, L.K. Hansen. *Neural Network for Modeling and Control of Dynamic Systems.* Springer-Verlag, London, 2000.

[15] M. Norgaard. *Neural Network Based System Identification Toolbox, Tech. Report. 00-E-891.* Department of Automation, Technical University of Denmark, 2000.

[16] A.J. Oberholster, P.S. Heyns. On-line fan blade damage detection using neural network. *Mechanical Systems and Signal Processing*, **20**: 78–93, 2006.

[17] J. Sanz, R. Perera, C. Huerta. Fault diagnosis of rotating machinery based on auto-associative neural networks and wavelet transforms. *Journal of Sound and Vibration*, **302**: 981–999, 2007.