www.czasopisma.pan.pl PAN www.journals.pan.pl
INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2011, VOL. 57, NO. 3, PP. 401–406
Manuscript received July 6, 2011; revised August 2011. DOI: 10.2478/v10177-011-0055-2

Recurrent Neural Networks for Predictive Maintenance of Mill Fan Systems

Petia D. Koprinkova-Hristova, Mincho B. Hadjiski, Lyubka A. Doukovska, and Simeon V. Beloreshki

***Abstract*—In the present paper we focus on online monitoring system for predictive maintenance based on sensor automated inputs. Our subject was a device from Maritsa East 2 power plant – a mill fan. The main sensor information we have access to is based on the vibration of the nearest to the mill rotor bearing block. Our aim was to create a (nonlinear) model able to predict on time possible changes in vibrations tendencies that can be early signal for system work deterioration. For that purpose, we compared two types of recurrent neural networks: historical Elman architecture and a recently developed kind of RNN named Echo stet networks (ESN). The preliminary investigations showed better approximation and faster training abilities of ESN in comparison to the Elman network. Direction of future work will be increasing of predications time horizon and inclusion of our predictor at lower level of a complex predictive maintenance system.**

***Keywords*—Technical diagnosis, Thermal Power Plant (TPP), Recurrent Neural Networks (RNN), Distributed Control System (DCS), predictive maintenance.**

I. Introduction

MAIN source of damages and accidents in technological processes are erosion, corrosion, vibration, and depositions. The particularly illustrative example in metallurgical objects is wearing of ceramic insulation in contact with liquid metal. The degree of degradation of the facility is essential for determining future operational behavior. Therefore, the construction of diagnostic models is one of most important task. A major scientific problem is finding an approach for integration of diagnostic with analytical and based on current data models. Given the sharp worsening in the forecast quality of models according on the horizon of performance, finding a compromise between short-and long-term diagnostic and forecasting will be a problem. Here the use of training methods and evolutionary algorithms is inevitable. In the modern processes, control theory the predictive maintenance and operational impacts should be considered as two interrelated actions.

Fault detection is recognizing that a problem has occurred, even if you don't yet know the root cause. Faults may be detected by a variety of quantitative or qualitative means. This includes many of the multivariable, model-based approaches. It also includes simple, traditional techniques for single variables, such as alarms based on high, low, or deviation limits for process variables or rates of change, Statistical Process Control (SPC) measures, and summary alarms generated by subsystems.

The fault does not have to be the result of a complete failure of a piece of equipment, or even involve specific hardware. For instance, a problem might be defined as non-optimal operation or off-spec product. In a process plant, root causes of non-optimal operation might be hardware failures, but problems might also be caused by poor choice of operating targets, poor feedstock quality, poor controller tuning, and partial loss of catalyst activity, buildup of coke, low steam system pressure, sensor calibration errors, or human error.

Other elements of Operations Management Automation related to diagnosis include the associated system and user interfaces, and workflow (procedural) support to for the overall process. Workflow steps that might be manual or automated include notifications, online instructions, escalation procedures if problems are ignored, fault mitigation actions (what to do while waiting for repairs), direct corrective actions, and steps to return to normal once repairs are complete.

Automated fault detection and diagnosis depends heavily on input from sensors or derived measures of performance. In many applications, such as those in the process industries, sensor failures are among the most common equipment failures. Therefore, a major focus in those industries has to be on recognizing sensor problems as well as process problems. Distinguishing between sensor problems and process problems is a major issue. Our usage of the term "sensors" includes process monitoring instrumentation for flow, level, pressure, temperature, power, vibration and so on. In other fields such as network and systems management, it can include other measures such as error rates, CPU utilization, queue lengths, dropped calls, and so on. In addition, diagnosis as a decision support activity rather than a fully automated operation is common in the management of large, complex operations such as those found in the process industries and network and systems management.

Here we focus mainly on online monitoring systems, based on sensor or other automated inputs. In the presented paper, we have chosen to analyze a device from Maritsa East 2 power plant – a mill fan. The main sensor information we have access to is based on the vibration of the nearest to the mill rotor bearing block. Our aim was to create a (nonlinear) model able to predict on time possible changes in vibrations tendencies that can be early signal for system work deterioration.

This work has been supported by the Ministry of Education, Youth and Science under the Project BG051PO001-3.3.04/40/28.08.2009 and by the National Science Fund of Bulgaria under the Project no. DVU-10-0267/10.

P.D. Koprinkova-Hristova is with Institute of System Engineering and Robotics, Bulgarian Academy of Sciences (e-mail: pkoprinkova@icsr.bas.bg).

M.B. Hadjiski is with University of Chemical Technology and Metallurgy – Sofia, Bulgaria (e-mail: hadjiski@uctm.edu).

L.A. Doukovska is with Institute of Information and Communication Technologies, Bulgarian Academy of Sciences (e-mail: doukovska@iit.bas.bg).

S.V. Beloreshki is with Honeywell EOOD, Bulgaria (e-mail: simeon.belorechki@honeywell.com).

Artificial Neural Networks (ANN) is created initially to mimic human brain [1] but they are also recognized as universal approximations for any kind of non-linear dependences. In [2] it is proposed to use feed-forward ANN structure for mapping of vibration amplitude given time steps in the future accounting for its values for several time steps in the past. Although such approach showed good approximation abilities it needs preliminary investigation of the needed data time steps in the past to be included in to the ANN input vector. Inclusion of feedback connections in ANN architecture called Recurrent Neural Network (RNN) allows accounting for past input state influence to the network current output. However, training of such complicated architecture is not that easy and fast in comparison with the simple feedforward architectures.

A known by far Elman network architecture [3] consisting of a fully recurrent layer and linear output layer is one of first examples of RNN. Initially proposed algorithms for its training is backpropagation or its modifications [1] accepting that backward connections are constants. Other possible training algorithms are Extended Kalman Filter method (EKF) or backpropagation through time (BPTT) [1] that are much more computationally demanding but allow backward connections training and hence are much more accurate.

A recently proposed ESN structure [4]–[6] incorporates a randomly generated dynamic reservoir and easy trainable output neurons usually with linear transfer functions. This architecture is very similar to the Elman architecture with only difference that the recurrent layer is not fully connected and its connections are randomly generated. However, the training approach for ESN is different and faster because once generated reservoir connections weights as well as the input weights are not subject of training. The only trainable weights are that of the output connection matrix. Since the readout from the reservoir usually is linear dependence, the training can be done off-line within a single step or by using Recursive Least Squares algorithm (RLS) in on-line mode [4]. In search of improving reservoir quality algorithm for initial adjustment of reservoir connections matrix was proposed in [7]. It is called intrinsic plasticity (IP) and is aimed at increasing the entropy of the reservoir neurons outputs thus stabilizing its dynamic behavior. Such combined IP-RLS training was already successfully applied in [8], [9]

In our previous work, we [10] applied ESN as a model for prediction of changes in vibrations tendencies of mill fan system. Here our aim is to compare this newly developed kind of RNN with historical Elman RNN architecture. The preliminary investigations showed better approximation and faster training abilities of ESN in comparison to the Elman network. Direction of future work will be increasing of predications time horizon and inclusion of our predictor at lower level of a complex predictive maintenance system.

II. Problem Formulation

Maritsa East 2 thermal power plant (TPP) has built up eight blocks – 4×175 MW and 4×210 MW. In historic plan in 1962, a decision has been taken for building up Maritsa East 2 TPP, and since 1970, the electro energy of least price cost for the country is produced in Maritsa East 2 TPP. In the end of 1995 (19 Dec.) 8th energy block (215 MW) has been connected in parallel to the energy system of the country by which the second stage of Maritsa East 2 TPP enlargement was completed. Achieved installed capacity is 1450 MW. This turns Maritsa East 2 TPP into the biggest thermal power plant in the Balkans. After following reconstructions and modernizations installed capacity at the moment reaches 1556 MW as in the end of 2009 (Dec. 24) block 6 was cut off for the purpose of modernization and increasing its capacity to 230. The Maritsa East 2 TPP being the largest thermal power plant on the Balkan Peninsula and the choice of the given power plant is not occasional.

Honeywell's Experion control system is installed on Units 1, 3 and 4 in Maritsa East 2 thermal power plant. Using standard engineering tools computer models of the controlled units are created and entered in real-time database. Each parameter is collected and shown in control system's database in real time. It can be shown in different formats and appearance depending of the user needs – operator, shift engineer, maintenance personal, management etc.

Experion Process Knowledge System (PKS) is a cost-effective open control and safety system that expands the role of distributed control. It addresses critical manufacturing objectives to facilitate sharing knowledge and managing workflow. Experion provides a safe, robust, scalable, plant-wide system with unprecedented connectivity through all levels of the plant as illustrated in the following high-level view of the architecture. The Experion unified architecture combines DCS functionality and a plant-wide infrastructure that unifies business, process, and asset management to Facilitate knowledge capture; Promote knowledge sharing; Optimize work processes; Accelerate improvement and innovation. The Experion platform is well suited for both small and large systems. It provides the power and flexibility required to handle the full spectrum of process control and safety applications.

Data gathering and collection of as much as possible data is the fundamental of decision support. It is not practical to measure all parameters in a typical power plant, so we can judge about them by other parameters and overall parameters trends. Data gathering for long period can give good indication how process was developed during the time – for example we can judge if overhaul improved or not equipment operation. In other words, system provides additional information for additional analysis. Depending on the significance and skills needed, conclusion can be made by operator, shift supervisor or other manager depending on duties and responsibilities.

Maybe most important fact is that during system refurbishment project, big part of knowledge and good practice are embedded in the system. In the past, operator needed to be highly skilled, nowadays part of their knowledge and best control practice are taken into consideration during design of the control algorithms. Direction is to have fully automated plant, although in most cases this is practically impossible. The idea is to reduce the human factor and related to this factor incidents. Besides all parameters, needed for the operators, additional data and calculations are available to upper level management for decisions that can have economical and financial impacts.

This is next level in control – information and process calculations, different and specific production costs etc., which are crucial for the manager. For example, let's imagine that unit is working at 150MW load and this is optimal mode of operation, but there is a demand for 170MW – management should make calculations if it is economically beneficial to provide this additional 20MW, but to put unit in less optimal operating point.

This is second level of information control system is providing – this information is more relevant not to the process control, but decision making. Manager could decide to reduce slightly efficiency of the unit, by putting it out of optimal operational point, but to benefit more from the additional price he is receiving for providing more power on demand. That's why it is so important that relevant information is available to all managers at different levels.

In the presented paper, we have chosen to analyze a device from Maritsa East 2 power plant – a mill fan. The mill-fans are used to mill, dry and feed the coal to the burners of the furnace chamber. They are together milling and transporting devices. Mill-fans are most often used for power plants burning brown and lignite coal. In general, these are large centrifugal fans which suck flue gases with temperature around 800-1000 degC from the top of the furnace chamber. In the same pipe, the coal is feed, thus diminishing the drying agent temperature and drying the coal prior entering the fan. The coal is being milled by the fast rotating rotor of the fan and turn into coal dust. This dust is transferred to separator that returns the bigger particles to the fan. The separator can be tuned for a desired dust granulometric size. One of the most important parameters to control is the discharge temperature of the dust-air mixture. For the considered mill-fan, it should be between 145-195 °C. Lower than 145 °C may cause clogging of the mill and higher than 195 °C may cause the dust to be fired in the ducts prior the burners. This temperature is also a measurement for the load of the mill. The lower the temperature the higher the load is – more coal is fed to the mill. The part, which suffers the most and needs care, is the rotor of the mill fan. Because of the abrasive effect of the coal, it wears out and should be repaired by welding to add more metal to the worn out blades.

The boiler which milling system is studied is a Benson type once-though sub-critical boiler. There are four mills per boiler. Each mill fan system has four radial bearings – two in the mill and two in the motor. The DCS installed on the site is Honeywell Experion R301 Process. All the data used in the present research are obtained from the historian system of the Distributed Control System (DCS).

On the next Figure, the structure scheme of the boiler as well as its main parameters is shown. These include: Steam output, t/h 690; Steam pressure, MPa (kgf/cm2) – at boiler outlet 15,4 (157) and at re-heater outlet 3,5 (35,7); Steam temperature, °C – primary steam temperature 540 and reheat steam temperature 540; Secondary steam flow, t/h 600; Design fuel: Lignite; Design gross efficiency of the boiler, % 89,9; Weight content of NOx in flue gas (at $\alpha = 1.4$), mg/Nm3 315.

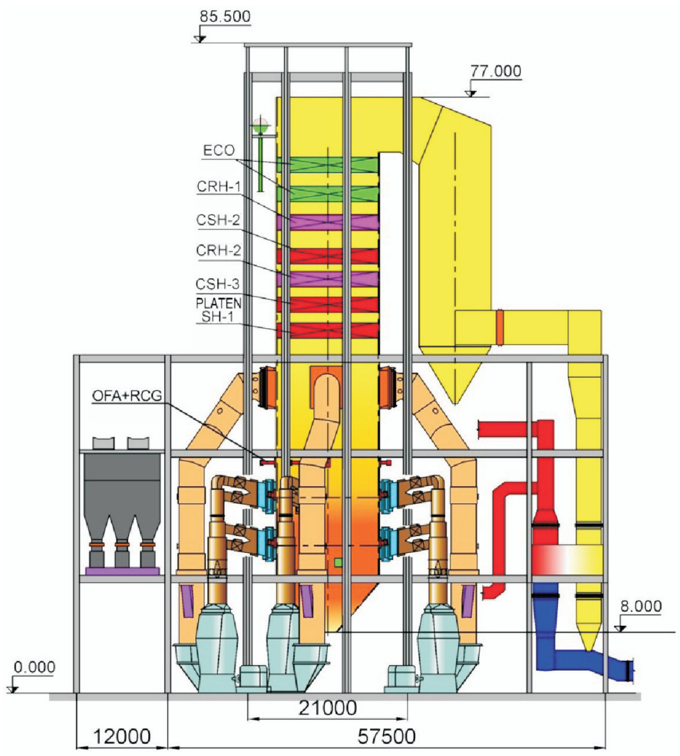

Fig. 1. Boiler Ep-690-15,4-540 LT.

Prognostic maintenance of a mill fan is considered in the presented paper. It is based on the vibration of the nearest to the mill rotor bearing block. The observation period is 16.12.2010 – 16.01.2011. On 31.12.2010, the rotor of the mill-fan is changed. After the replacement, it has been working for 378 hours. The period chosen allows for vibrations analyzes before and after the replacement. It is observed that after the replacement the vibrations with new rotor have higher amplitudes than with the worn out one. This is because of the abrasive wear out of the rotor – the blades become thinner and the rotor becomes lighter. The new rotor is heavier so the vibrations are more intensive even though the rotor has been carefully balanced.

In the present paper, we will describe briefly two considered RNN architectures and their training algorithms.

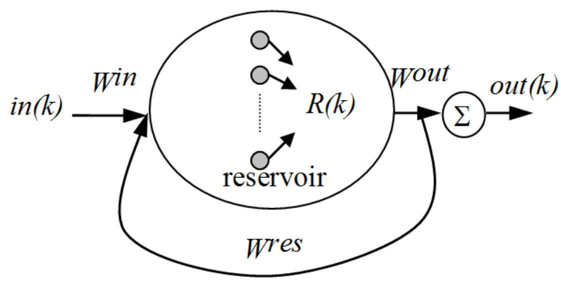

Fig. 2. Elman network.

A. Elman Network

The Elman network architecture is shown on Figure 2.

In order to make comparison with ESN later we adopt some of terminology typical for that architecture. The fully connected recurrent hidden layer is named "reservoir". All of reservoir or hidden neurons have connections from the input $in(k)$ at discrete time instant k and from reservoir output $R(k-1)$ one time step backwards. The reservoir output is nonlinear function f^{res} of its inputs linear combination as follows:

$$R(k) = f^{res}\Big(W^{in}in(k) + W^{res}R(k-1)\Big) \tag{1}$$

W^{in} and W^{res} are $n_R \times n_{in}$ and $n_R \times n_R$ matrices, where n_{out}, n_{in} and n_R are the sizes of the corresponding vectors out, in and R. Feedback connections are between every two neurons in the hidden layer, i.e. all the elements of matrix W^{res} are non-zero. The network output is linear combination of all hidden neurons outputs as follows:

$$out(k) = W^{out}R(k) \tag{2}$$

W^{out} is a $n_{out} \times (n_{in} + n_R)$ matrix.

Since training of feedback connections matrix W^{res} is complicated, one possibility is to fix them to be constants, all with values equal to 1. Thus subject of training are the rest of connections weights. BPTT-EKF algorithms [1] allow training of feedback connections too but at considerably higher computational cost and corresponding time needed.

In our investigation, we used EKF algorithm for training all the Elman network connections weights. For that purpose, we used the Matlab function developed by Yi Cao [11].

B. Echo State Network

ESNs are a kind of recurrent neural networks that arise from so called "reservoir computing approaches" [6].

The basic ESN structure is shown in Figure 3 below. Here the notions are the same as those for Elman network architecture. The main difference is that the reservoir neurons are randomly connected and some of the elements of W^{res} are zero. W^{in} and W^{res} are randomly generated and are not trainable. Another difference is that direct connection from input to output is also allowed. There are different approaches for reservoir parameter production [6]. A recent approach used in the present investigation is proposed in [7]. It is

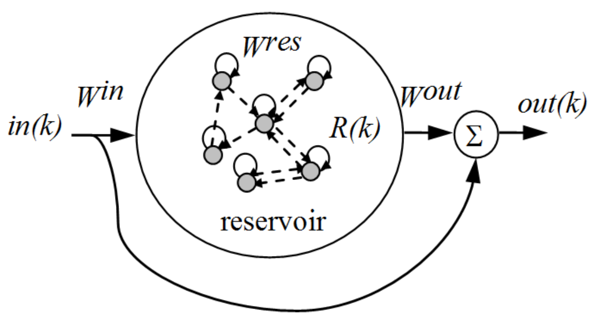

Fig. 3. Echo state network structure.

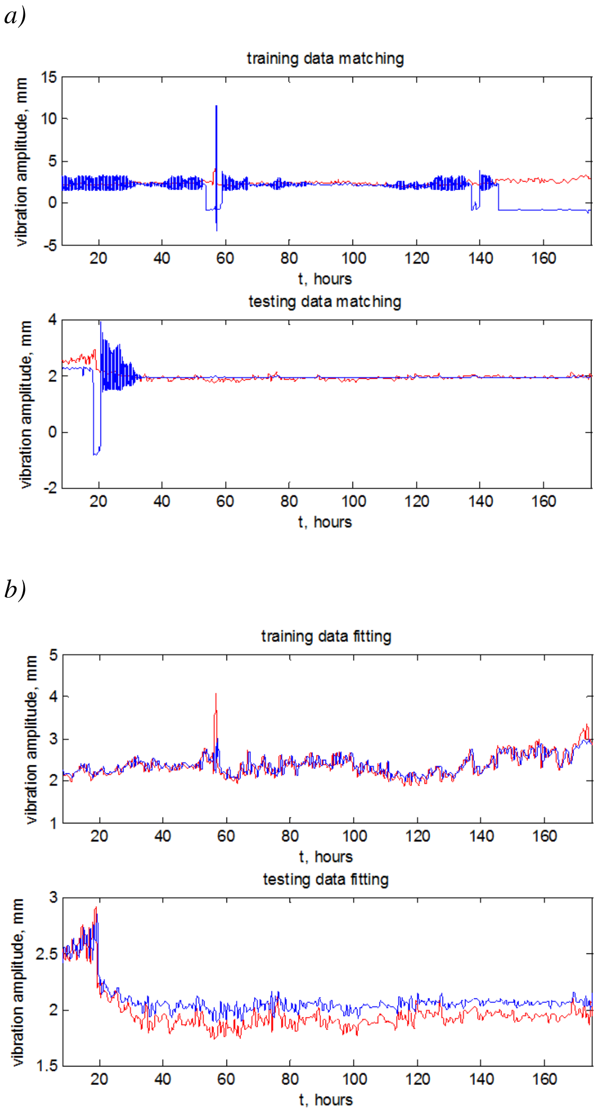

Fig. 4. a) Elman network training and testing with $\Delta t = 1$ minute, b) ESN network training and testing with $\Delta t = 1$ minute.

called intrinsic plasticity (IP). The algorithm suggests initial adjustment of reservoir matrix, aimed at increasing the entropy of the reservoir output. ESN training can be done in an off-line or an on-line mode. For on-line training, the RLS algorithm [6] was proposed. It is claimed that it converges fast and it is less computationally expensive in comparison to BPTT-EKF methods [12].

Here we used the free software from [13] with our addition of IP pre-training algorithm from [7].

III. Results and Discussion

Since in real industrial conditions there are many vibration sources. That is why wavelet de-noising method was applied [14]. In our case, it is done by soft heuristic threshold of the wavelet coefficients using symlets for wavelet decomposition at level 5. After filtration, the data are divided into two sets – for training and testing of both RNN structures.

The aim was to train RNN able to predict vibrations amplitude several time steps ahead using current measurement

TABLE I
TRAINING AND TESTING NRMSE AND TRAINING TIME IN CASE ΔT=1 MINUTE

NN structure	NRMSE Training	NRMSE Testing	Training time
Elman	4.8215	3.2841	more than 10 hours
ESN	0.51571	0.72461	about 20 minutes

as input. For that purpose our RNNs has one input and one output – for the vibration amplitude at current and future time respectively. Here we've trained several Elman and ESNs to predict vibration amplitude 30 minutes ahead. In both cases, the reservoir contains only 10 neurons and initial spectral radius for ESN was 0.9.

First, we tried to train RNN predictors using measurement with one minute time step. The results for training and testing data sets are shown on Figure 4 for the Elman and ESN structures respectively. The red line presents real measurements data while the blue – RNN predictions in both cases.

Table I summarizes the Normalized Root-Mean-Square Error (NRMSE) for both cases and approximate training time needed. The RLS method definitely showed that needs much less time to train much more perfectly ESN. In contrast, the Elman network is unable to be well trained and demonstrates unstable behavior. This can be explained also with huge amount of used data – more than 19000 items for training and testing data sets.

Next, we tried to decrease the data set size using measurement at every 10 minutes. Thus, the number of training and testing data sets was 10 times decreased. Figure 5 below presents the data fitting results for the training and testing data sets for Elman and ESN structures again. The red line presents real measurements data while the blue – RNN output predictions in both cases.

Table II summarizes the NRMSE for both cases and approximate training time needed. Again, the ESN is better trained. In that case, Elman network was trained with considerably good accuracy too due to lower amount of data.

IV. CONCLUSIONS

In the present investigations, we've tested two similar RNN architectures – well known Elman network and newly developed Echo state network – as well as their training algorithms

TABLE II
TRAINING AND TESTING NRMSE AND TRAINING TIME IN CASE ΔT=10 MINUTE

NN structure	NRMSE Training	NRMSE Testing	Training time
Elman	0.52098	0.69633	more than 5 hours
ESN	0.48746	0.68533	about 10 minutes

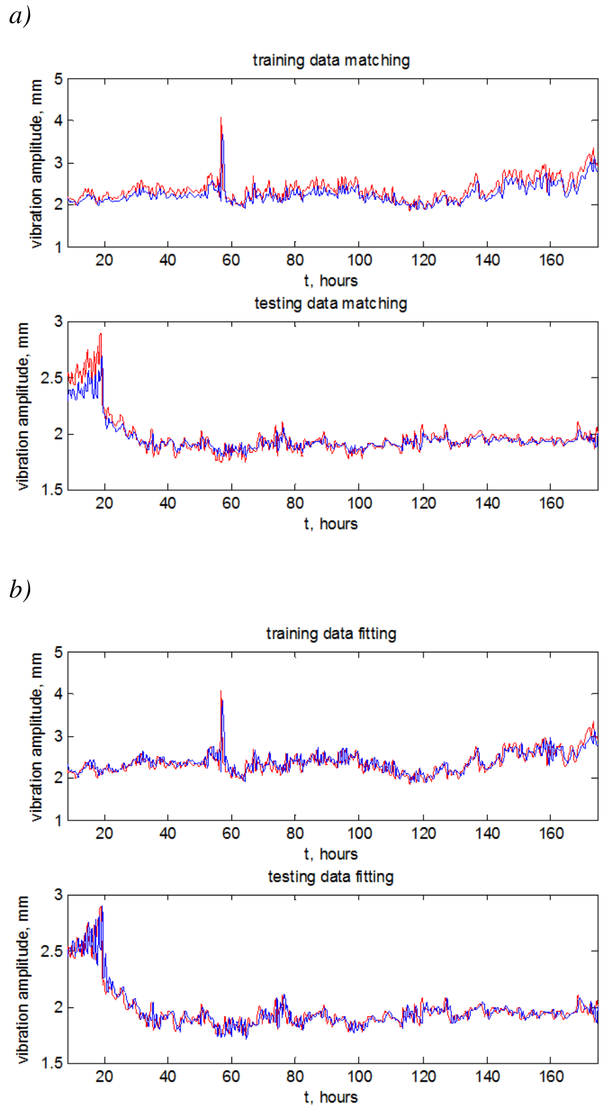

Fig. 5. a) Elman network training and testing with $\Delta t = 10$ minute, b) ESN network training and testing with $\Delta t = 10$ minute.

(EKF and RLS). Both RNN structures were trained to predict vibrations amplitude of a mill fan system. The obtained results demonstrated the superiority of ESN structure with respect to the data fitting accuracy and training time needed. The obtained results are encouraging. Our next aim will be to improve the ESN predictions quality for bigger time horizon by increasing the reservoir size or using of global feedback connection within ESN architecture.

REFERENCES

[1] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. Prentice Hall International, Inc., 1999.
[2] F. E. Ciarapica and G. Giacchetta, "Managing the Condition-Based Mintenance of a Combined-Cycle Power Plant: An Approach Using Soft Computing Techniques," *Journal of Loss Prevention in the Process Industries*, vol. 19, pp. 316–325, 2006.
[3] J. L. Elman, "Finding Structure in Time," *Cognitive Science*, vol. 14, pp. 179–211, 1990.
[4] H. Jaeger, "Tutorial on Training Recurrent Nural Networks, Covering BPPT, RTRL, EKF and the "Echo State Network"Approach," German National Research Center for Information Technology, GMD Report 159, 2002.

[5] ——, "Adaptive Nonlinear System Identification with Echo State Networks," in *Advances in Neural Information Processing Systems 15*. Cambridge, MA: MIT Press, 2003, pp. 593–600.
[6] M. Lukosevicius and H. Jaeger, "Reservoir Computing Approaches to Recurrent Neural Network Training," *Computer Science Review*, vol. 3, pp. 127–149, 2009.
[7] B. Schrauwen, M. Wandermann, D. Verstraeten, and J. J. Steil, "Improving Reservoirs Using Intrinsic Plasticity," *Neurocomputing*, vol. 71, pp. 1159–1171, 2008.
[8] P. Koprinkova-Hristova and G. Palm, "Adaptive Critic Design with ESN Critic for Bioprocess Optimization," *Lecture Notes in Computer Science*, vol. 6353, pp. 438–447, 2010.
[9] P. Koprinkova-Hristova, M. Oubbati, and G. Palm, "Adaptive Critic Design with Echo State Network," in *2010 IEEE Intenrational Conference on Systems, Man and Cybernetics SMC'2010*, Istanbul, Turkey, October 10-13 2010, pp. 1010–1015.
[10] T. Balabanov, P. Koprinkova-Hritstova, L. Doukovska, M. Hadjiski, and S. Beloreshki, "Neural Network Model of Mill-Fan System Elements Vibration for Predictive Maintenance," in *2011 Intenrational Symposium on Innovations in Intelligent SysTems and Applications (INISTA)*, Istanbul, Turkey, June 15-18 2011, pp. 410–414.
[11] Http://www.mathworks.com/matlabcentral/fileexchange/18289.
[12] D. Prokhorov, "Echo State Networks: Appeal and Challenges," in *Proceedings of International Joint Conference on Neural Networks (IJCNN)*, Montreal, Canada, 2005, pp. 1463–1466.
[13] Http://www.reservoir-computing.org/software: Simple and very simple Matlab toolbox for Echo State Networks by H. Jaeger and group members.
[14] D. L. Donoho, "De-Noising by Soft-Thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, 1995.