International Conference
AUTOMATICS AND INFORMATICS'2013
3-7 October 2013, Sofia, Bulgaria

LEVENBERG-MARCQUARDT TRAINING APPROACH FOR RECURRENT FUZZY-NEURAL NETWORK

Y. Todorov[1], M. Terziyska[2], M. Petrov[2]

[1]Institute of Information and Communication Technologies, Bulgarian Academy of Sciences, Sofia 1113, Acad. G. Bonchev str., bl. 2, tel. +359-2-979-29-28,
[2]Technical University - Sofia, branch Plovdiv, 4000, 25, Ts. Dustabanov st.,

yancho.todorov@iit.bas.bg , terzyiska@gmail.com , mpetrov@tu-plovdiv.bg

Abstract: This paper describes the development of a Levenberg-Marcquardt learning approach for the consequent part of the fuzzy rules in recurrent Takagi-Sugeno type inference. The recurrent relation in the proposed fuzzy-neural network represents a global feedback from the fuzzy-neural network output to its relevant inputs, being fuzzified in the next training sample. To prove the efficiency of the proposed fuzzy-neural structure, simulation experiments for prediction of Mackey-Glass chaotic time series are performed. A comparison with classical Gradient descent method is also studied.

Key words: Takagi-Sugeno, Fuzzy-Neural networks, Recurrent Neural networks, Levenberg-Marcqardt, Gradient descent

INTRODUCTION

In general, dynamic systems are complex and nonlinear. An important step in nonlinear control is the development of a nonlinear model. The problem of identification consists of choosing an identification model and adjusting the parameters such that the response of the model approximates the response of the real system to the same input.

In recent years, computational intelligence techniques, such as neural networks, fuzzy logic and combined neuro-fuzzy systems algorithms have become very effective tools for identification and control of nonlinear plants. For this purpose, a general fuzzy-neural approach based on multiple LTI models around various function points has been proposed. The so-called, Takagi-Sugeno (TS) approach is a convex polytopic representation, which can be obtained either through mathematical transformation or through achieved linearization around various operating points. The main advantage of the TS model is the soft transition through any operating regions [1-2].

A recurrent neural network (RNN) is a network of neurons with feedback connections. It can learn many behaviors: sequence processing tasks, algorithms, programs that are not learnable by traditional machine learning methods. This explains the rapidly growing interest in *artificial* RNN implementations for technical applications: for instance to map input sequences to output sequences, with or without a teacher. Such structures are computationally more powerful and more plausible than other adaptive approaches such as: feedforward networks. Their recent applications include adaptive robotics and control, attentive vision, protein analysis, stock market prediction, and many other sequence problems [3]. During the last decades, RNNs are widely investigated and many models have been developed [4-7].

Many approaches also show different suitable applications of TS inference realized as recurrent fuzzy-neural network. (RFNNs) implementations are proved to be suitable in describing complicate dynamical systems than and FNNs, because they can handle the time-varying inputs or outputs through its own natural temporal operation. RFNNs have one or more feedback loops that allow them to capture the dynamic response of a system. Due to its dynamic characteristic and relatively simple architecture, the RFNNs are useful tools for most real-time applications, such as modeling and predictive control. For instance, in [8] is suggested a recurrent TS network for sliding mode control. In [9] a recurrent neuro-fuzzy system for implementation of long-range prediction fuzzy model has been proposed. In this recurrent neuro-fuzzy model the network output is fed back to the network input through one or more time delay units. Levenberg-Marquardt algorithm with regularization is used for adjusting crisp weights and biases of the feed-forward and feed-back connections of the recurrent neuro-fuzzy network. The suggested neuro-fuzzy network is applied to modeling and control of a neutralization process. As well, many different recurrent neuro-fuzzy networks are reported in [10-13]. Stable predictive controller based on RFNN model is described in [14]. In [15-16] are presented predictive controllers based on RFNN where the feedback is after the fuzzyfication layer. A novel Type-2 RFNNs with asymmetric membership functions are proposed in [17-18]. TSK-type RFNNs for modeling and control are described in [19]. These RFNNs have additional inputs which increase the number of the fuzzy rules and the computational burden, respectively.

This paper describes the development of a Recurrent Takagi-Sugeno (RTS) type fuzzy-neural network with global feedback and hybrid learning algorithm. The proposed learning approach is based on a Gradient descent algorithm for adjusting the fuzzy rules premise parameters and Levenberg-Marcquardt algorithm to adjust the rules consequents in order to improve the model qualities and to minimize the possible model oscillations. The efficiency of the proposed modeling approach is tested in prediction of a Mackey-Glass chaotic time series and shows a better model performance compared to classical Gradient descent approach.

RECURRENT FUZZY-NEURAL NETWORK

Since the middle of the 1980s, TS fuzzy-neural models have attracted a great deal of attention from industrial practitioners and academic researchers, especially because they can effectively approximate a wide class of nonlinear systems. Thus, in discrete time by using the NARX representation model (*Nonlinear AutoregRessive model with eXogenous inputs*) the proposed recurrent fuzzy-neural network can be derived as:

$$y(k) = f_y(x(k)) \tag{1}$$

where the unknown nonlinear function f_y can be approximated by Takagi-Sugeno type fuzzy rules:

$$R^{(i)} : if\ x_1\ is\ \tilde{A}_1^{(i)}\ and\ x_p\ is\ \tilde{A}_p^{(i)}\ then\ \ f_y^{(i)}(k) \tag{2}$$

$$\begin{aligned} f_y^{(i)}(k) = {} & a_1^{(i)} y_m(k-1) + a_2^{(i)} y_m(k-2) + \ldots + a_{ny}^{(i)} y_m(k-n_y) + \\ & + b_1^{(i)} u(k) + b_2^{(i)} u(k-1) + \ldots + b_{nu}^{(i)} u(k-n_u) + b_0^{(i)} \end{aligned} \tag{3}$$

where $(i)=1,2,\ldots N$ denotes the number of the fuzzy rules $R^{(i)}$. A_i is an activated fuzzy set defined in the universe of discourse of the input x_i and the crisp coefficients $a_1, a_2, \ldots a_{ny}, b_1, b_2, \ldots b_{nu}$ are the coefficients into the Sugeno function f_y. The input vector $\boldsymbol{x}$ contains regressors in notion the input/output history dependence. On Fig.1 is shown the schematic diagram of the proposed recurrent TS fuzzy-neural network.

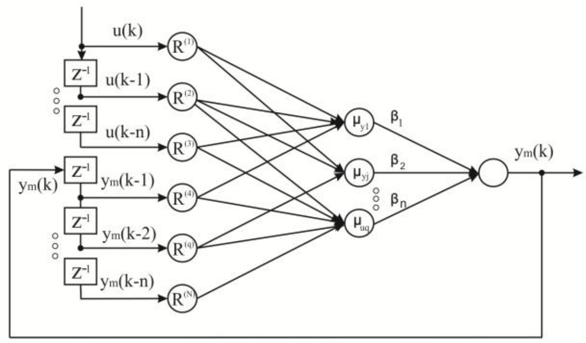

Fig.1. Schematic diagram of the proposed recurrent fuzzy-neural network

The identification of the proposed recurrent network requires the two main groups of unknown parameters to be determined: the number of membership functions, their shape and the parameters of the function f_y in the consequent part of the rules. For this purpose, in this work a simplified fuzzy-neural approach is applied [20-21].

Learning algorithm for designed recurrent TS fuzzy-neural network. A two step learning procedure based on minimization of an instant error measurement function $E=\varepsilon^2/2$ and $\varepsilon(k)=y(k)-y_m(k)$ between the process output and the model output, is implemented. During the learning process, two groups of parameters in the fuzzy-neural architecture – premise and consequent parameters are under adaptation. The consequent parameters are the coefficients $a_1, a_2, \ldots a_{ny}, b_1, b_2, \ldots b_{nu}$ in the Sugeno function f_y and they are calculated by the following equations when using Gradient descent method:

$$\begin{aligned} \beta_{ij}(k+1) &= \beta_{ij}(k) + \eta(y - y_M)\bar{\mu}_y{}^{(j)}(k) x_i(k), \\ \beta_{0j}(k+1) &= \beta_{0j}(k) + \eta(y - y_M)\bar{\mu}_y{}^{(j)}(k) \end{aligned} \tag{4}$$

where η is the learning rate and β_{ij} is an adjustable i^{-th} coefficient (a_i or b_i) in the Sugeno function f_y of the j^{-th} activated rule. The premise parameters are the centre c_{ij} and the deviation σ_{ij} of a Gaussian fuzzy set defined as:

$$m_{ij}^{(i)}\left(x_i\right) = exp\left(-\left(x_i - c_i\right)/\,2s_i\right)^2 \tag{5}$$

The parameters of a fuzzy set are calculated by using the following equations:

$$c_{ij}(k+1) = c_{ij}(k) + \eta(y - y_M)\bar{\mu}_y{}^{(j)}(k)[f_y^{(i)} - y_m(k)]\frac{[x_i(k) - c_{ij}(k)]}{c_{ij}^2(k)} \tag{6}$$

$$\sigma_{ij}(k+1) = \sigma_{ij}(k) + \eta(y - y_M)\bar{\mu}_y{}^{(j)}(k)[f_y^{(i)} - y_m(k)]\frac{[x_i(k) - \sigma_{ij}(k)]^3}{\sigma_{ij}^2(k)} \tag{7}$$

Levenberg-Marcquardt approach as training method for hybrid learning of the network parameters. To improve the efficiency of the proposed recurrent fuzzy-neural network, a Levenberg-Marcquardt (LM) approach for adjusting the rules consequent parameters, is applied. Since the LM method requires the computation of the second order derivative of the defined error cost term, it can be rewritten:

$$\Delta\beta = -\left[\nabla^2 E\left(\beta\right)\right]^{-1} \nabla E\left(\beta\right) \tag{8}$$

The Hessian and the Gradient of $E(\beta)$ are expressed as:

$$\nabla E\left(\beta\right) = J^T\left(\beta\right) e\left(k\right) \tag{9}$$

$$\nabla^2 E\left(\beta\right) = J^T\left(\beta\right) J\left(\beta\right) + \sum_{j=1}^{N} e_j\left(k\right) \nabla^2 e_j\left(k\right) \tag{10}$$

where the dimension of the *Jacobian* matrix is *(NxN$_p$)*; *N*- the adjustable parameters in the network. As LM is a Newton type approach the second term in (10) is assumed equal to zero. Therefore, the update rule, according to (8) became*:*

$$\nabla\beta = -\left[J^T\left(\beta\right) J\left(\beta\right) + \lambda I\right] J^T\left(\beta\right) e\left(k\right) \tag{11}$$

where λ is the LM parameter and I is identity matrix. The *Jacobian* according to adjustable parameters is calculated as:

$$J^T\left(\beta_{0j}\right) = \bar{\mu}_y^{(j)},\ J\left(\beta_{0j}\right) = \left[\ \bar{\mu}_y^{(j)}\ \right]^T \tag{12}$$

$$J^T\left(\beta_{i,j}\right) = \ \bar{\mu}_y^{(j)} x_i\left(k\right),\ J\left(\beta_{i,j}\right) = \left[\ \bar{\mu}_y^{(j)} x_i\left(k\right)\right]^T \tag{13}$$

Finally, the recurrent equations for the rule consequent parameters are derived as follows:

$$\begin{aligned} \beta_{0j}(k+1) = {} & \beta_{0j}(k) + \\ & + \eta\left[\bar{\mu}_y{}^{(j)}\left(\bar{\mu}_y{}^{(j)}\right)^T + \lambda(k) I_i\right]^{-1} (y - y_M)\bar{\mu}_y{}^{(j)} \end{aligned} \tag{14}$$

$$\begin{aligned} \beta_{ij}(k+1) = {} & \beta_{ij}(k) + \\ & + \eta\left[\bar{\mu}_y{}^{(j)} x_i(k)\left(\bar{\mu}_y{}^{(j)} x_i(k)\right)^T + \lambda(k) I_i\right]^{-1} (y - y_M)\bar{\mu}_y{}^{(j)} \end{aligned} \tag{15}$$

The proposed LM approach is coupled in a hybrid learning algorithm for the assumed recurrent fuzzy-neural network, where the rules premise parameters are scheduled at each step by using the Gradient descent method, while for training of the rules consequents, the LM method is employed.

EXPERIMENTAL MODEL EVALUATION

Mackey-Glass chaotic time series. Chaos is a common dynamical phenomenon in various fields [22] and different definitions as series representations exist. Chaotic time series are inherently nonlinear, sensitive to initial conditions and difficult to be predicted. For that purpose, the chaotic time series prediction based on measurement is a practical technique for studying characteristics of complicated dynamics [23] and evaluation of the accuracy of different types of nonlinear models as RNN's. In this study a Mackey-Glass [24] chaotic time series is used to evaluate the performance of the proposed fuzzy-neural network.

The Mackey-Glass (MG) time series are defined as:

$$\frac{dq}{dt} = b\frac{q_r}{1 - q_t{}^n} - gq, \qquad \text{where } b, g, n > 0 \tag{16}$$

where β, γ, τ, n are real numbers, and q_τ represents the value of the variable q at time $(t-\tau)$. Depending on the values of the parameters, this equation displays a range of periodic and chaotic dynamics. The following parameters of the MG time series are assumed: β=0.3, γ=0.1, n=1, τ=10.

The generated chaotic time series are used as input to second order differential equation model of a continuous stirred tank reactor (CSTR) in order to simulate a toughest example of nonlinear dynamics which has to be predicted by the proposed fuzzy-neural network. The dynamic equations of the nonlinear plant are given by:

$$\dot{x}_1 = -x_1 + D_a(1-x_1)\exp\left(\frac{x_2}{1+x_2/\phi}\right) \quad (17)$$

$$\dot{x}_2 = -(1+\delta)x_2 + BD_a(1-x_1)\exp\left(\frac{x_2}{1+x_2/\phi}\right) + \delta u \quad (18)$$

where x_1 and x_2 represent the dimensionless reactant concentration and the reactor temperature, respectively. The control action u is dimensionless cooling jacket temperature. The physical parameters in the CSTR model equations are D_a, ϕ, B and δ which correspond to the Damkhler number, the activated energy, the heat of reaction and the heat transfer coefficient, respectively. Based on the nominal values of the system parameters, D_a=0.072, ϕ=20, B=8 and δ=0.69, the open-loop CSTR exhibits three steady states $(x_1,x_2)_A$=(0.144, 0.886), $(x_1,x_2)_B$=(0.445, 2.75) and $(x_1,x_2)_C$=(0.765, 4.705), where the upper and the lower steady states are stable, whereas the middle one is unstable. The control objective here is to bring the nonlinear CSTR from the stable equilibrium point $(x_1,x_2)_A$ to the unstable one $(x_1,x_2)_B$ [25].

The simulation experiments are performed under equal initial conditions for the model parameters as well as, for the parameters of the chaotic series. Many switching criteria are proposed for use with the LM method. In the presented application, it is used the well known notation: when the error term is increased, compared to its previous value the λ is multiplied by some factor λ_{inc} when the error term is increased and to some other λ_{dec}. when it is decreased. The values of the scaling factors are: λ_{inc}=5 and λ_{dec}=0.05.

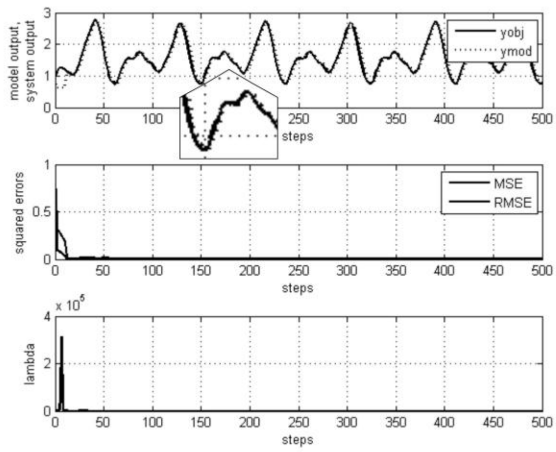

Fig.2. Levenberg-Marcqurdt learning approach with λ_{in}=4.000

On Fig.2 and Fig.3 are demonstrated the obtained results with two different initial values of the LM parameter λ of 4.000 and 40.000. There are shown the predicted signal and its actual value, as well as, the transient responses of the MSE, RMSE and the parameter λ. The simulation experiments show expected model behaviour. In the very first steps of the learning algorithm, when the model error is relatively high, the LM approach switches to higher values of the parameter λ, in order to diminish the modelling error and switches back to smaller values when the error term is relatively small.

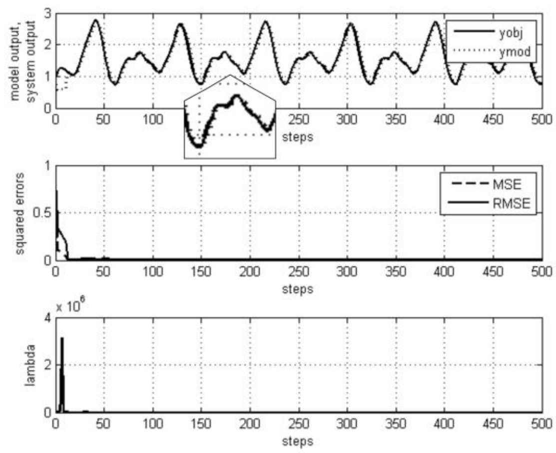

Fig.3. Levenberg-Marcqurdt learning approach with λ_{in}=40.000

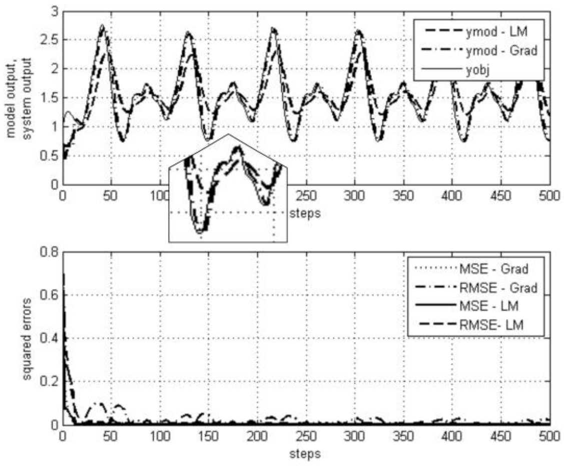

Fig.4. Comparison between the Gradient descent and Levenberg-Marcquardt approach

On Fig.4 is demonstrated a comparison between the classical Gradient descent approach for learning the rules consequents in the fuzzy-neural network and the proposed hybrid approach using the LM method. As can be observed from the obtained results, the LM method ensures a better model performance with small instant values of the modelling error.

In Table1 are compared the instant values of the RMSE's for the three investigated cases, measured in terms of 50 training cycles. As can be seen, the use of LM approach with higher values of λ_{in}, ensures the smallest possible prediction error, compared to other two approaches, which made it the best possible solution.

The obtained results show also that, when the parameter λ is small, the method represents a quadratic approximation and when it is large, the Hessian is negligible and the LM method works similarly as Gradient descent algorithm. At first iterations, LM works as a gradient method and as it gets near the optimal point it gradually switches to Newton based approach. When LM parameter gets smaller, LM finds a locally linear solution, precisely and quickly.

Table.1

Step	LM λ_{in}=4.000	LM λ_{in} =40.000	Gradient descent
	RMSE	RMSE	RMSE
50	0.01743	0.01570	0.03650
100	0.00145	0.00110	0.00953
150	0.00077	0.00094	0.04300
200	0.00397	0.00360	0.00224
250	0.00158	0.00155	0.01070
300	0.00436	0.00043	0.02600
350	0.00039	0.00062	0.00870
400	0.00586	0.00565	0.12600
450	0.00105	0.00087	0.00463
500	0.00023	0.00003	0.02360

CONCLUSION

It was presented in this paper a hybrid approach for training a recurrent Takagi-Sugeno type neural-network. The fuzzy rules premise parameters are scheduled at each sampling period by using the Gradient descent method, while the rules consequents are calculated by employing the Levenberg-Marcquardt method. The performed investigations, for prediction of a generated Mackey-Glass chaotic time series, shows a better performance of the adopted approach, compared to classical Gradient method, for both type of fuzzy rules parameters. A crucial point is the selection of the initial value of the LM parameter λ and switching thresholds. The use of higher values of λ ensures a fast convergence of the method, as well as minimal instant values of the predicted error.

REFERENCES

1. Mendes, J.. Adaptive fuzzy generalized predictive control based on Discrete-Time T-S fuzzy model. In Proceedings of the IEEE Conference of Emerging Technologies and Factory Automation, 2010, pp. 1-8.
2. Chadli M., P. Borne. Multiple Models Approach in Automation: Takagi-Sugeno Fuzzy Systems, Wiley, 2012.
3. Schmidhuber, J., F. Gomez, S. Fernandez, A. Graves, S.Hochreiter. Sequence Learning with Artificial Recurrent Neural Networks. *Cambridge University Press*, in preparation. http://www.idsia.ch/~juergen/rnnbook.html
4. Li, X., Z. Chen, Z. Yuan. Simple recurrent neural network-based adaptive predictive control for nonlinear systems. Asian Journal of Control, 4(2), 2002, pp. 231-239.
5. Rankovic, V., I. Ž. Nikolić. Neural Network Model Predictive Control of Nonlinear Systems Using Genetic Algorithms, International Journal of Computational Communication, 7(3), 2012, pp. 540-549.
6. Seyab, R. Al, Y. Cao. Differential Recurrent Neural Network based Predictive Control. Computers and Chemical Engineering, 32(7), 2008, pp. 1533-1545.
7. Mohajerin, N.. Identification and Predictive Control Using Recurrent Neural Networks, Studies from the Department of Technology at Örebro University, Sweden, 2012.
8. Lin, F., P. Huang. Recurrent fuzzy neural network controller design using sliding-modecontrol for linear synchronous motor drive. IEE Proceedings of Control Theory and Applications, 151(4), 2004, pp. 407 – 416.
9. Wu, S., M.J Er. Dynamic fuzzy neural networks- a novel approach to functionapproximation. IEEE Transactions Systems Man Cybernetics-Prat B Cybernetics , 30(2), 2000, pp. 358-364.
10. Lin, C., C. Chen, Y. Lee. Recurrent fuzzy-neural network adaptive hybrid control for linearized multivariable systems. Journal of Intelligent
11. Juang, C.F., K.C. Ku. A recurrent network for fuzzy temporal sequence processing and gesture recognition. IEEE Transations On Systems, Man, and Cybernetics, Part. B: Cybernetics, 35(4), 2005, pp. 646-658.
12. Juang C.F.. A TSK type recurrent fuzzy network for dynamic systems processing by neural network and genetic algorithms. IEEE Trans. on Fuzzy Systems, 10(2), 2002, pp. 155-170.
13. Tamura, H., K. Tanno, H. Tanaka, C. Vairappan, Z. Tang. Recurrent Type ANFIS using Local Search Technique for Time Series Prediction. In Proceedings of Asia Pacific Conference on Circuits and Systems, 2008, pp. 380-383.
14. Chi-Huang, L., L. Chi-Ming. Stable predictive control based on recurrent fuzzy neural networks, Proceeding on 8th Asian Control Conference (ASCC), 2011, pp. 897-901.
15. Chi-Huang, L., T. Ching-Chih. Generalized predictive control using recurrent fuzzy neural networks for industrial processes, Journal of Process Control, 17, 2007, pp.83-92.
and Fuzzy Systems, 17(5), 2006, pp.479-491.
16. Mendes, J., N. Sousa, R. Araujo. Adaptive predictive control with recurrent fuzzy neural network for industrial processes, 16th IEEE ETFA Conference, 2011, pp. 1-8.
17. Ching-Hung, L., HH. Chang, C. Ting Kuo, J. Chieh Chien, T. Wei Hu. A Novel Recurrent Interval Type-2 Fuzzy neural Network for Nonlinear Channel Equilization, Proceeding of the Int. MultiConf. of Eng. and Computer Sci., 2009, pp. 7-12.
18. Chia-Feng J., L. Chun-Feng, C. Po-Han. Dynamic Plant Control Using Recurrent Fuzzy Controller with Ant Colony Optimization in Real Space, In Proc. of the Int. IEEE Conf. on Systems Man and Cybernetics, 2010, pp. 1134-1138.
19. Chia-Feng, J., L. Yang-Yin, C. I-Fang. Dynamic System Identification Using A Type-2 Recurrent Fuzzy Neural Network, In. Proc. of the 7th Asian Control Conference, 2009.
20. M. Terzyiska, Y. Todorov, A. Mitev, M. Petrov. Nonlinear Model Based Predictive Controller using a Fuzzy-Neural Hammerstein model. In Proc. of the Int. Conf. ‘Modern Trends in Control’, 2006, pp. 299-308.
21. M. Terzyiska, Y. Todorov, M. Petrov. Nonlinear Model Predictive Controller with Adaptive Learning rate Scheduling of an internal model. In Proc. of the Int. Conf. ‘Modern Trends in Control’, 2006, pp. 289-298.
22. Yao, J., J. Mao, W. Zhang. Application of Fuzzy Tree on Chaotic Time Series Prediction, IEEE Proc. of Int. Conf. on Aut. and Logistics, 2008, pp.326-330.
23. Lai Y.. Recent developments in chaotic time series analysis. International Journal of Bifurcation and Chaos, 13(6), 2003, pp.1383-1422.
24. Diaconescu, E.. The use of NARX Neural Networks to predict Chaotic Time Series. WSEAS Transactions on Ccomputer Rresearch, 3(3), 2008, pp. 182-191.
25. Ray, W.H.. Advanced Proceess Control, McGraw-Hill, New York, 1981.