Nonlinear System Identification based on NARX Network

Hongwei Liu
School of Aerospace Science
Beijing Institute of Technology
Beijing, China
1132520084@qq.com

Xiaodong Song
School of Aerospace Science
Beijing Institute of Technology
Beijing, China
xd_song@bit.edu.cn

***Abstract*—This paper discusses identification of nonlinear system with nonlinear AutoRegressive models with eXogenous inputs (NARX). NARX network is a dynamic neural network which appears effective in the input-output identification of both linear and nonlinear systems. When identifying them by NARX model, the first step is to collect training data and the final results vary considerably with different training data. The paper compares the training results of three kinds of signals, including SPHS signal, Gaussian white noise and mixed signal. Our results show the response characteristics of NARX model trained by different signals can be used to design the input training signal.**

Keywords—NARX network; system identification; SPHS signal; Gaussian white noise

I. INTRODUCTION

For complex systems, it is hard to obtain systematical precise mathematical model, and even harder to get parametric model. So identification of nonlinear systems is one of the problems we have to face when designing a controller. In recent year, because of rapid development of neural network theory, more and more variety of neural network is used in system identification. Nerandra demonstrated that artificial neural networks could be used successfully for the identification and control of nonlinear dynamic systems[1]. Some research works have adopted neural network to study the system parameter estimation. Annaswamy and Yu developed the Θ-Adaptive NN block and NN recursive estimation methods for parameter estimation of a nonlinear dynamic system[2]. The ARAM models by means of evolutionary algorithms are used for model identification and parameter estimation[3]. However, the defect of adopting parameter identification is needing model's structure, so identification based on input-output data is necessary. Neural network identification is a dynamic system by using foregoing input and output. Sfak identifies universal motor dynamically with neural network[4]. Sastry identifies and controls dynamical systems with a kind of neural network with memory ability[5]. Coban put forward a kind of context layered locally recurrent neural network (CLLRNN) for dynamic system identification[6]. Koriani have applied recurrent neural network to real-time identification and control[7]. Among these techniques reported, a new kind of neural network based on nonlinear AutoRegressive models with eXogenous inputs (NARX models) [8, 9] has been used to identify complex systems.

In the process of the NARX model for system identification, the first step obtains the experimental data that describes the underlying intrinsic features of the nonlinear system. Sahoo has adopted Gaussian white noise with SNR 30dB[10,11]. Phm Huy anh applies Pseudo Random Binary Signal (PRBS) to identify a pneumatic artificial muscle (PAM) robot arm[12]. Salami takes PRBS signal to identify the DC motor driven rotary motion system[13]. Tijani use a sinusoidal input of varying frequency to excite the small scale unmanned helicopter while keeping the system at the desired hovering operating point as much as possible[14]. Safak have applied the Schroeder-phased harmonic sequence (SPHS) signal as input signal to identify DC motor[4]. Sastry proposes mixture of two signals, two-thirds of them take iid sequence uniform and the rest takes Sinusoidal Signal[5]. Annabestani makes more signals mixing. They use a square signal, a chirp signal, a sinusoidal signal, a PRBS signal and a sampled Gaussian noise signal, and take them to identify system of Ionic Polymer Metal Composite (IPMC) respectively[15]. However, there are not enough research about the influence caused by different input signal. So the paper will discuss the problem and show the comparison results.

The rest of the paper is organized as follows. Section II presents an overview of the NARX model. Three kinds of signal is proposed to train the network in Section III. Experiment and result are given in Section IV. Finally, the paper is concluded in Section V.

II. NARX NETRORK OVERVIEW

NARX network describes discrete nonlinear system by using past input and output data[1]. To the SISO system, the expression is shown below:

$$y(k) = \varphi[y(k-1), y(k-2), \ldots, y(k-n); u(k), u(k-1), \ldots, u(k-m)] \quad (1)$$

where $u(k)$ and $y(k)$ represent input and output values, respectively, of the system at time step k, n and m represent input and output orders, and φ represents nonlinear mapping function.

978-1-4799-7862-5/15/$31.00 © 2015 IEEE

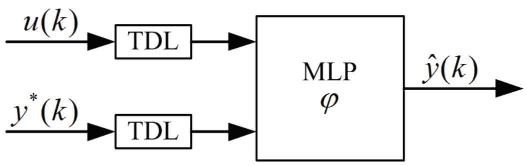

Fig. 1. Architecture of NARX network.

In the paper, nonlinear mapping function φ is achieved by multilayer perceptron (MLP) network. In order to make notation more clearly, now the regulations are made: $\hat{y}(k)$ represents NARX network estimated or predicted value and $y^*(k)$ represents output's feedback into NARX network model. The NARX model is as shown in Fig. 1.

When $y^*(k) = \hat{y}(k)$, the network is called NARX parallel network. On the other hand, when $y^*(k) = y(k)$, it is called serial-parallel architecture. In the process of identifying with NARX model, serial-parallel structure is usually adopted that has two advantages as below. First, it can make feedback signal more accurate and improve the success rate of identification. Second, serial-parallel structure can easily be converted to parallel structure to proceed system simulation. Therefore, the paper trains with serial-parallel structure and analyzes precision of NARX identification model with parallel structure.

In order to show NARX structural model visually, the general architecture of the NARX network with three input's tapped delays and four output's tapped delays is shown in Fig. 2. In Fig. 2, z^{-1} represents delay operator. Multilayer perceptron network acts as the nonlinear system, and tansig transfer function is used in the hidden layer while purelin function is used in output layer. This structure can be extended to other network dimensions. Usually, the parameters of the total NARX network includes tapped delays of input and output, the number of hidden layers and neuron nodes in each of the layers. The parameters are usually obtained through heuristic process or optimization algorithm. Levenberg-Marquardt training algorithm is used to adjust network weight. As mentioned in [16], the Levenberg-Marquardt algorithm was found to be the fastest training algorithm, but requiring more memory with the same error convergence bound compared to the other algorithms. According to above introduction, the procedure of system identification with NARX model consists of five steps:

- Step 1: collect the input and target data.
- Step 2: define the input and output tapped delays.
- Step 3: initialize the structure of the NARX network.
- Step 4: train the NARX in serial-parallel architecture.
- Step 5: simulate the NARX parallel by the test signal.
- Step 6: calculate the mean squared error.

III. Input Signal Selection

The basis of NARX identification is input-output data. There are two requirements. The first is that input-output data must be sufficient to acquire the system's behavior. The second is that the volume of data cannot be too big resulting in over fitting. Hence, it is a problem that must be solved firstly that how to design an effective input signal. The paper puts forward several training input signal.

A. SPHS signal

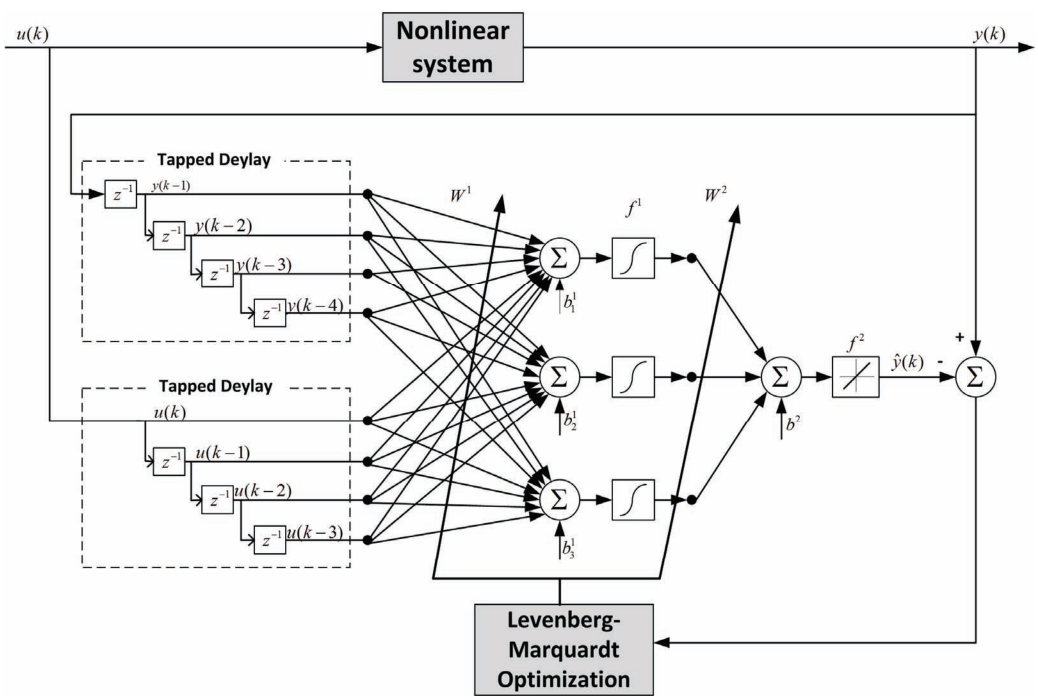

Fig. 2. Architecture of NARX network.

The first signal is Schroeder-phased harmonic sequence (SPHS) signal. Schroeder [17] has developed a method of synthesizing such a signal with any desired spectrum by adding a series of waves with different phases using the equation:

$$u(t)=\sum_{k=1}^{NH} a_k \cos(\frac{2\pi k}{T_p}+\theta_k) \tag{2}$$

where NH is the number of harmonics, a_k is the amplitude and θ_k is the angle of the kth harmonic, and T_p is the fundamental period of the SPHS signal. According to the above form, SPHS signal is composited with several signals that has certain spectrum. In practical application, a flat-spectrum signal ($a_1=a_2=a_3=\cdots=a_k=\text{const}$.)with a desired bandwidth can be formed using the formulae[4].

$$u(t)=\sum_{k=1}^{NH}\sqrt{\frac{2}{NH}}\cos(\frac{2\pi k}{T_p}+\theta_k) \tag{3}$$

$$\theta_k=\frac{2\pi}{NH}\sum_{i=1}^{k} i \tag{4}$$

where the bandwidth equals $2\pi NH/T_p$ rad/s, the angular frequency interval equals $2\pi/T_p$.

B. Gaussian White noise signal

The perfect shite noise has infinite width and power. It is nonexistent in real world. Therefore, it is considered as white noise approximately that band-limited signal whose average power is equally distributed in actual use.

C. Mixed signal

From the neural network learning process, different kinds of signals should be input into network, which contributes to improve network's generalization ability and identify a system comprehensively. It can also increase learning areas and identification ability of NARX network. The common signals are square wave signal, chirp signal, sinusoidal signal, PRBS signal, Gaussian white noise, and so on. Therefore, the third signal is designed as shown in Fig. 3 according to above signals. The first part is chirp signal, the second part is Gaussian white noise and the third part is PRBS signal.

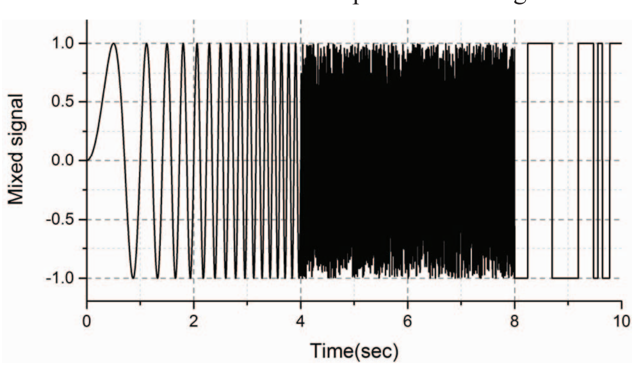

Fig. 3. Mixed signal.

IV. EXPERIMENT AND RESULT

In this section, the linear model and nonlinear model are discussed respectively with NARX model. At the same time, it is analyzed the influence of system identification results with different input signals. In the paper, training and testing performances are determined by the root-mean-squared error (RMSE) criteria:

$$RMSE=\sqrt{\frac{1}{K}\sum_{k=1}^{K}(y(k)-\hat{y}(k))^2} \tag{5}$$

where K represent the size of data volume.

There are three kinds of training signals used in this section. The first is SPHS signal, the second is Gaussian white noise and the third is mixed signal. For the testing part, the input sequence consisting of mixtures of sinusoids and constant signal is used:

$$u(t)=\begin{cases}\sin(5\pi t), & t\le 1\text{s}\\ 1.0, & 1.0\text{s}<t\le 1.5\text{s}\\ -1, & 1.5\text{s}<t\le 2.0\text{s}\\ 0, & 2.0\text{s}<t\le 2.5\text{s}\\ 0.3\sin(5\pi t)+0.6\sin(3\pi t) & \\ +0.1\sin(12\pi t), & 2.5\text{s}<t\le 5\text{s}\end{cases} \tag{6}$$

A. Example 1

To the first example, a nonlinear identification problem is used to checkout performance of NARX network. In this example, the nonlinear dynamic system to be identified is described by the following difference equation [1, 6]:

$$y(k)=\frac{y(k-1)y(k-2)y(k-3)u(k-2)[y(k-3)-1]+u(k-1)}{1+y^2(k-2)+y^2(k-3)} \tag{7}$$

To the nonlinear model, the input $u(k)$ is needed to be two tapped delays and the output $y(k)$ is needed to be the three tapped delays. The number of initialized hidden layer of network is 1 and the number of the first hidden layer neurons is 10. The training signals are three kinds of signals.

1) SPHS 信号

A signal consisting of three SPHS periods, each covering the amplitude range of 0.5, has been synthesized for training the NARX network. This method produced slightly better experimental results than applying s single SPHS. The training signal is constructed as:

$$u(t)=\begin{cases}0.5\text{SPHS}(NH,T_p)-0.5 & t<3\text{s}\\ 0.5\text{SPHS}(NH,T_p) & 3\text{s}\le t<6\text{s}\\ 0.5\text{SPHS}(NH,T_p)+0.5 & 6\text{s}\le t<10\text{s}\end{cases} \tag{8}$$

where $NH=20$, $T_p=10\text{s}$ [4]. The input and output of the system are displayed in Fig. 4 and Fig. 5. Fig. 5 shows the output of the dynamic system and the NARX network for the test input given by equation (6). The NARX network has achieved the RMSEs of 0.0037 and 0.0385 for training and testing. The prediction error between them is plotted in Fig. 5

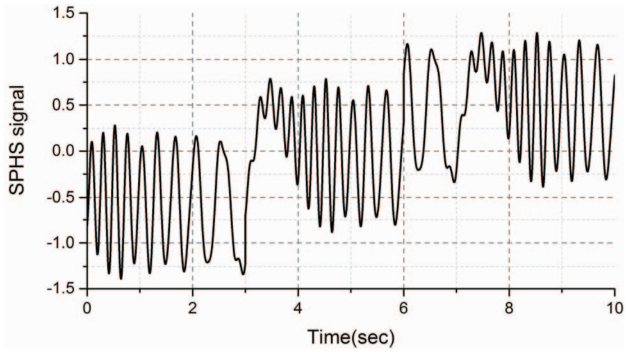

Fig. 4. SPHS signal.

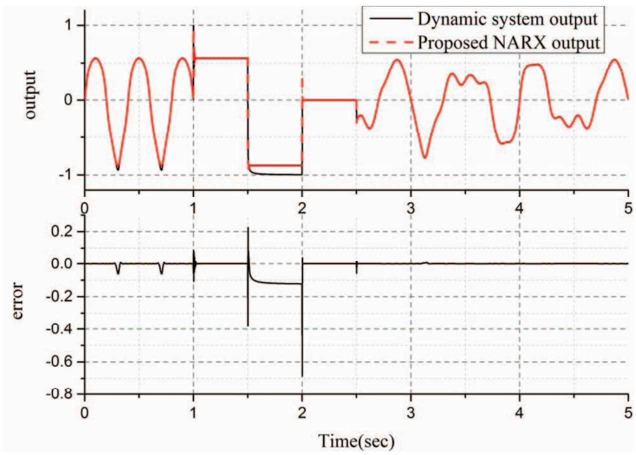

Fig. 5. The performance of the NARX network trained by SPHS signal.

and it is easy to find that the identification perform well except the step signal.

2) Gaussian white noise

Train with Gaussian white noise. Fig. 6 shows the output of the dynamic system and the NARX network for the test input

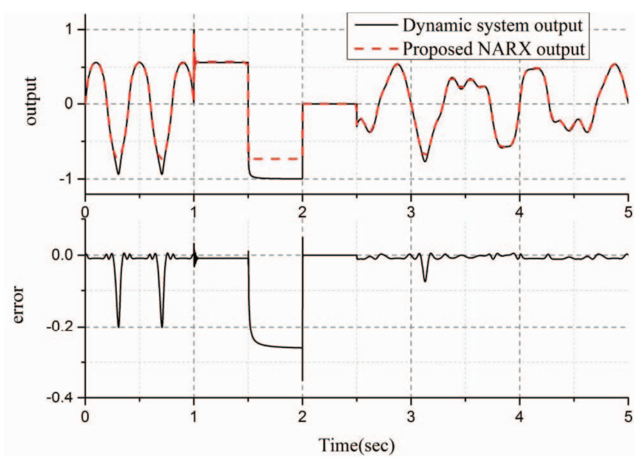

Fig. 6. The performance of the NARX network trained by Gaussian white noise.

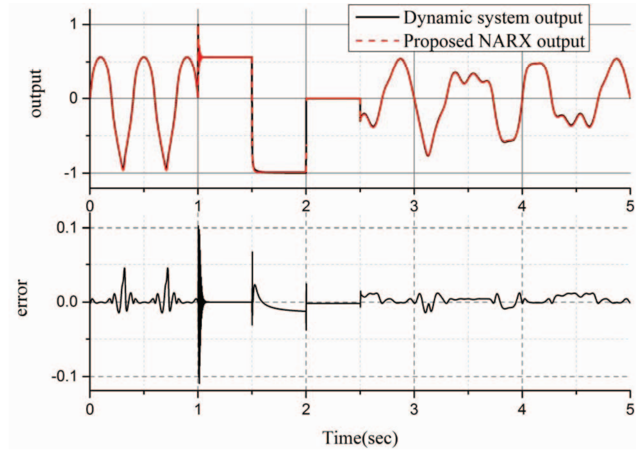

Fig. 7. The performance of the NARX network trained by mixed signal.

given by equation (6). The NARX network has achieved the RMSEs of 0.0249 and 0.0826 for training and testing. The prediction error between them is also plotted in Fig. 6, it is observed that NARX model trained with Gaussian noise presents better in the step response and worse in the sine.

3) Mixed signal

Train with mixed signal. Fig. 7 shows the output of the dynamic system and the NARX network for the test input given by equation (6). The NARX network has achieved the RMSEs of 0.0072 and 0.0085 for training and testing. The prediction error between them is plotted in Fig. 7, it is observed that NARX model presents better trained with mixed signal.

4) Comparison

The result of three parts is shown in Table I. From the Table I, it is seen that mixed signal presents best, Gaussian white noise presents worst and the error of SPHS signal in identifying step signal is large while close in the step signal. To show the advantage of the NARX network, it is compared with the ERNN, MNN and CLRNN [6], the details are given in Table II. It can be observed that using NARX model can get better results trained by mixed signal showing lower RMSE both in the training and testing phase.

TABLE I. PERFORMANCE COMPARISON FOR NONLINEAR SYSTEM IDENTIFICATION BASED ON DIFFERENT TRAINING SIGNAL

Train signal type	**SPHS signal**	**Gaussian white noise**	**Mixed signal**
RMSE for training	0.0037	0.0249	0.0072
RMSE for testing	0.0385	0.0826	0.0085
Training time steps	117	113	258

TABLE II. DIFFERENT TYPES OF MODELS' PERFORMANCE COMPARISON

Network structure	**MNN**	**ERNN**	**CLLRNN**	**NARX**
RMSE for training	-	0.039	0.016	0.0072
RMSE for testing	0.075	0.041	0.020	0.0085

B. Example 2

It is discussed that identification effect of SPHS signal, Gaussian white noise and mixed signal by a typical nonlinear model in Example 1. Then it is found that neural network can get better effect combined training by mixed signals. In this example, the conclusion will be certificated by three models.

The first model is a second-order linear system described by[6]:

$$\begin{aligned} y(k) &= 1.8398533y(k-1) - 0.8607080y(k-2) \\ &\quad + 0.0106881u(k-1) + 0.0101666u(k-2) \end{aligned} \tag{9}$$

The second model expression is:

$$y(k) = \frac{1-e^{-g(k)}}{1+e^{-g(k)}} \tag{10}$$

where $g(k)$ is a function of:

$$\begin{aligned} g(k) &= 2.2y(k-1) + 0.5y(k-2) \\ &\quad + u(k) - 1.25u(k-1) + 0.65u(k-2) \end{aligned} \tag{11}$$

The third system is:

$$\begin{aligned} y(k) &= 0.15y^2(k-1) + 0.3y(k-2) \\ &\quad + 0.6u^3(k) + 0.18u^2(k-1) - 0.24u(k-2) \end{aligned} \tag{12}$$

By using the three kinds of Training signal and test signal (6), result is shown in Fig. 8 and Fig. 9. It can be seen from the graph that RMSE for testing can reach least by mixed signal and the identification effect is the best. At all, when the NARX model is used to identify, the training effect is better by signal that is mixed by different signals than by single signal. At the same time, it can be seen that random signal can increase the accuracy of response to jump signal of NARX model, while continuous signal can increase the accuracy of response to gradient signal of NARX model.

V. Conclusion

In the process of actual controlling, it is especially important to identify nonlinear system. The paper achieved to identify nonlinear model with NARX model using the structure of serial-parallel connection. The MLP is constructed with past input and output samples, which is trained by Levenberg-Marquardt algorithm. Among it, the design of training signal is

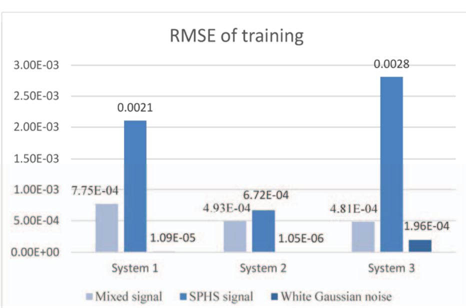

Fig. 8. The RMSE of training in Example 2.

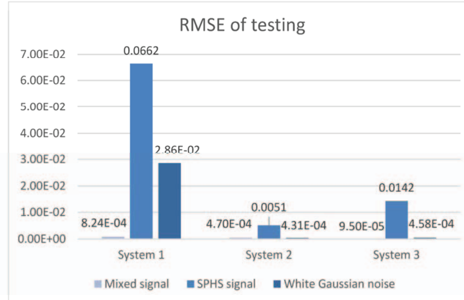

Fig. 9. The RMSE of testing in Example 2.

discussed specially. According to comparison of SPHS signal, Gaussian white signal and mixed signal, it is found that model trained by mixed signal has the best effect and it solves the design problem of training signal in process of identifying with NARX model.

References

[1] K. S. Narendra and K. Parthasarathy, "Identification and control of dynamical systems using neural networks," IEEE Transactions on Neural Networks, vol. 1, pp. 4-27, 1990.

[2] A. M. Annaswamy and S. Yu, "-adaptive neural networks: a new approach to parameter estimation," IEEE Transactions on Neural Networks, vol. 7, pp. 907-918, 1996-01-01 1996.

[3] S. Rolf, J. Sprave and W. Urfer, "Model identification and parameter estimation of ARMA models by means of evolutionary algorithms,", New York, NY, USA, 1997, pp. 237-43.

[4] K. K. Safak and O. S. Turkay, "Experimental identification of universal motor dynamics using neural networks," Mechatronics, vol. 10, pp. 881-96, 2000-12-01 2000.

[5] P. S. Sastry, G. Santharam and K. P. Unnikrishnan, "Memory neuron networks for identification and control of dynamical systems," IEEE Transactions on Neural Networks, vol. 5, pp. 306-319, 1994.

[6] R. Coban, "A context layered locally recurrent neural network for dynamic system identification," Engineering Applications of Artificial Intelligence, vol. 26, pp. 241-50, 2013-01-01 2013.

[7] M. M. Korjani, O. Bazzaz and M. B. Menhaj, "Real time identification and control of dynamic systems using recurrent neural networks," Artificial Intelligence Review, vol. 30, pp. 1-17, 2008-01-01 2008.

[8] M. M. Korjani, O. Bazzaz and M. B. Menhaj, "Real time identification and control of dynamic systems using recurrent neural networks," Artificial Intelligence Review, vol. 30, pp. 1-17, 2008-01-01 2008.

[9] M. V. Kumar, S. N. Omkar, R. Ganguli, P. Sampath, and S. Suresh, "Identification of helicopter dynamics using recurrent neural networks and flight data," Journal of the American Helicopter Society, vol. 51, pp. 164-174, 2006-01-01 2006.

[10] H. K. Sahoo, P. K. Dash and N. P. Rath, "NARX model based nonlinear dynamic system identification using low complexity neural networks and robust H filter," Applied Soft Computing, vol. 13, pp. 3324-34, 2013-07-01 2013.

[11] K. Nishiyama, "Robust estimation of a single complex sinusoid in white noise-H filtering approach," IEEE Transactions on Signal Processing, vol. 47, pp. 2853-6, 1999-10-01 1999.

[12] H. P. H. Anh and K. K. Ahn, "Hybrid control of a pneumatic artificial muscle (PAM) robot arm using an inverse NARX fuzzy model," Engineering Applications of Artificial Intelligence, vol. 24, pp. 697-716, 2011-01-01 2011.

[13] M. J. E. Salami, I. B. Tijani, A. I. Abdullateef, and M. A. Aibinu, "Hybrid algorithm for NARX network parameters' determination using differential evolution and genetic algorithm,", UK, 2013, pp. 012062 (10 pp.).

[14] I. B. Tijani, R. Akmeliawati, A. Legowo, and A. Budiyono, "Nonlinear identification of a small scale unmanned helicopter using optimized NARX network with multiobjective differential evolution," Engineering Applications of Artificial Intelligence, vol. 33, pp. 99-115, 2014.

[15] M. Annabestani and N. Naghavi, "Nonlinear identification of IPMC actuators based on ANFIS-NARX paradigm," Sensors and Actuators A (Physical), vol. 209, pp. 140-8, 2014-03-01 2014.

[16] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the Marquardt algorithm," IEEE Transactions on Neural Networks, vol. 5, pp. 989-93, 1994-11-01 1994.

[17] M. R. Schroeder, "Synthesis of low-peak-factor signals and binary sequences with low autocorrelation," IEEE Transactions on Information Theory, vol. IT-16, pp. 85-9, 1970-01-01 1970.