مجلة القادسية للعلوم الصرفة المجلد (12) العدد (4) لسنة 2007

System Identification Using a Hybrid Feed Forward Neural Networks

م.م علي حسين حسن
جامعة ذي قار /كلية العلوم
م.م رشيد بشير
جامعة ذي قار /كلية العلوم

الخلاصة:–

في هذا البحث استعملت شبكة عصبية ذات الانتشار الأمامي تحوي ثلاث طبقات من العصبونات(neurons) وبطريقة الانتشار العكسي كطريقة للتعلم وذلك لتكوين منظومة تعريف الأنظمة. وتم بناء الطبقة الوسطية (المخفية) للشبكة المقدمة على شكل هجين بدالتي تنشيط مختلفتين وهما الدالة الخطية ودالة (المماس لدالة القطع الزائد). بينت النتائج العملية قدرة الشبكة المقترحة الكبيرة على تعريف مختلف الأنظمة الخطية وغير الخطية وبسرعة تعلم كبيرة نسبيا"وبدون الحاجة إلى استخدام شبكات ذات دوال تنشيط مناظرة للمنظومة المراد تعريفها.

Abstract:-

A back-propagation feed forward neural network with three layers was applied for different systems identifications. The intermediate layer (hidden layer) of the proposed neural network is built with hybrid activation functions provided with two different activation functions namely linear activation function and hyperbolic tangent activation function. The results show the ability of the network to identify linear as well as nonlinear system with high speed of convergence comparing with traditional neural network structures. Without the need to build networks that have activation factions correspond the applied system.

1. Introduction

System Identification is the process of finding a model that best produces a data, obtained by a system with a known input. Obtaining a model of a system is quite useful in studying its behaviour. This model can be obtained by either the use of physical laws that govern the system or by identification procedures [1] (which can be performed by processing input / output data obtained by performing various experiments). Once a good model is obtained, it can be used for the analysis of the system properties, prediction and controller design.

There are two types of identification categories:-

(1) Parametric approaches such as least square, recursive least square, generalized least square, maximum likelihood an instrumental variables.

(2) Non-Parametric approaches which include transient analysis, frequency analysis, correlation analysis, spectrum analysis and Neural Networks (NNs).

In general non-parametric methods are easy to use but they give moderately accurate models. In neural network approach, the model is obtained by adjusting a number of connection weights arranged in a sufficient layers, such that the difference between the system's output and the neural network's output is minimized.

A neural network is a mathematical model, which is normally used to identify dynamic systems. Its benefit is the capability to identify a system when the model structure is not defined.

There are different kinds of neural networks (all of them are characterized by a specific architecture or some other features) and many different training algorithms used for identification problems for example Scott and Harmon[2,3,4], Codeca and Casella[5], Li[6], Gil et al[7], Ko et al[8], Kosmatopoulos et al[9], Alanis et al[10],

2. Neural Networks:-

Artificial Neural Networks (ANN's), also know as connectionist models, are rapidly evolving facet of artificial intelligence (AI) [2]. In particular, the main idea is to reproduce the intelligence and the capability to learn from examples, simulating the brain neuronal structure on a calculator. ANN's are interconnected networks of simple elements which interact with the objects of the real world in the same way as biological nervous system does. Because of the biological basis for the artificial neural networks, it is not surprising that many of the terms used in their study are borrowed from neurophysiology [2]. The processing units are neurons, nodes or processors; while the connections between these units are known as interconnected, synapses or weights. The pattern of the connections between the units determines the architecture of the network, which in the extremes, can be fully interconnected (recurrent neural networks) or connected in one direction only (feed forward neural networks).

ANN's must be learned before its using. There are many learning algorithms that can be used to train ANN's; back-propagation (with its modifications) is currently the mainstay of artificial neural networks learning and it selected to train the proposed neural network.

2.1. Hybrid Feed Forward Neural Networks:-

The feed forward neural network is the most used neural network architecture, it based on series connection of neuron layers, each one composed by a set of neurons connected in parallel. The signals flow the input layer through the hidden layer(s) to the output layer via uni-directional connections. There are one or more layers between the input and output layers.

The proposed neural network as shown in figure 1, consists of a three layers which are input layer with four neurons, the hidden layer with six neurons halves of them have a linear activation functions and the other halves have a hyperbolic tangent activation function, and the output layer has one output neuron.

By assuming that the sum of inputs to the i^{th}-cell of the K^{th}-layer is S_i^k, its output is O_i^k, the weighting value from i^{th}-cell of $(K-1)^{th}$-layer to the i^{th}-cell of the K^{th}-layer is wij. F (.) is the relation input between input and output (activation function). Therefore,

$$S_i^k = \sum_j w_{ij} S_i^{k-1} \quad \text{.......(1)}$$

$$O_i^k = F\,(S_i^k) \quad \text{.......(2)}$$

For the hidden layer, the activation function is

$$F(S_i^k)=\begin{cases} C.S_i^k & \text{for i} = 1,2,3 \\ \dfrac{\exp(S_i^k)-\exp(-S_i^k)}{\exp(S_i^k)+\exp(-S_i^k)} & \text{for i} = 4,5,6 \end{cases}$$

…………………..………………(3)

Where C is constant.

To improve the network precision the error back-probagation with momentum is introduced. That is the weighting values Neural Networks wij could be changed according to the error between the actual output and the desired output, so as to minimize the error.

$$\boldsymbol{E = \frac{1}{2}\sum_{i} (Oi - Yi)2} \quad ………………………………………………………………...(4)$$

2.2. Establish the Neural Network Model

In this work a parallel structure is used. The model Neural Network and the System (Plant) receive the same external inputs as shown in figure 2. The system and the network are two independent processes which share the same external inputs, but their outputs do not interfere with one another. The errors between these outputs are used to adjust the connection weights of the neural network.

3. Results

This section shows some simulation results. The training data excite all modes of the given plants, where the following input sequence is used for these purposes:-

$U_k=random[-3,3]$ *for* $1.0 \leq k \leq 500$ …..…………………..……………………(5)

The performance evaluation of the neural network is the root mean square (RMS) criteria used:-

$$RMS=\sqrt{\frac{\sum_k \ (y_p(k)-y_{net}(k))}{500}^{2}} \quad …………………………………….…………………...(6)$$

While the testing signal of 200 input samples used is as shown in figure 3:-

$$UK=\begin{cases} 0...\text{for } .k \leq 0 \\ 1.0..\text{for } \ 0 < k \leq 100.0 \ \ (7) \\ 1.0 + 0.5*\sin((k-70.)/4.0) + 0.05*\sin((k-70.)/8.0)..................\text{for } \ 100.0 < k \leq 200.0 \end{cases}$$

The simulation is made on the following three plants that differ in their degree of nonlinearity [1].

Plant Number 1:- Which is a linear in both input and output behaviour, as describe in equation (8).

$$y_k=A_1y_{k-1} + A_2y_{k-2} + B_1U_{k-1} + B_2\ U_{k-2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(8)$$

Where $A_1=1.752821,\ A_2= -0.818731,\ B_1=0.011698,\ B_2=0.010942$

The learning rate is adopted as 0.025, momentum constant is 0.35. and after 1000 iterations the results of RMS error is 0.001668 as shown in figure 4 and the testing signals is shown in figure 5.

Plant Number 2:- Which is a linear in the input and nonlinear in the output behaviour, as describe in equation (9).

$$y_k=A_1y_{k-1} + A_2y_{k-2} + A_3y^3_{k-2} + B_1\ U_{k-2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(9)$$

Where $A_1=1.04,\ A_2= -0.824,\ A3=0.130667,\ B_1= -0.16$

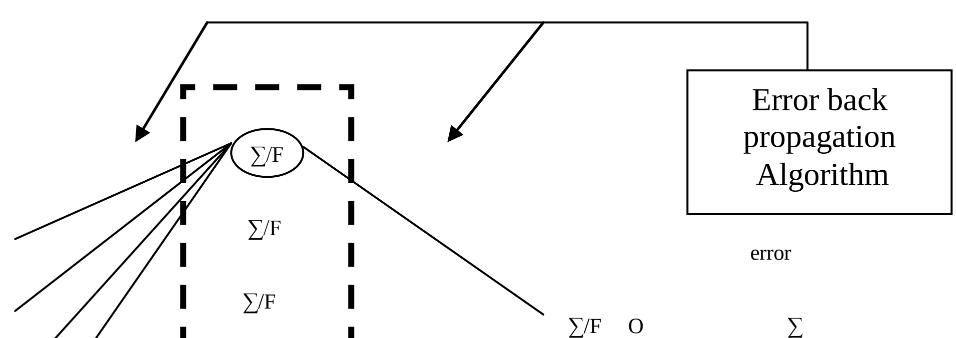
The learning rate is adopted as 0.025, momentum constant is 0.35 and after 1000 iterations the results of RMS error is 0.007278 as shown in figure 6 and the testing signals is shown in figure 7.

Plant Number 3:- Which is a strong nonlinear in both input and in the output behaviours, as describe in equation (10).

$$y_k= y_{k-1} /(1 + y^2_{k-1})+\ U^3_{k-1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(10)$$

The learning rate is adopted as 0.025, momentum constant is 0.45 and after 1000 iterations the results of RMS error is 0.001276 as shown in figure 8 and the testing signals is shown in figure 9.

The simulation show that the proposed network has a good identification results for the difference systems dynamics. This is because of the combination behaviour of the activation function where the linear part of the network give an excellent results for the linear characteristics of the plant while the nonlinear part of the network give an excellent results for the nonlinear characteristics of the plant.



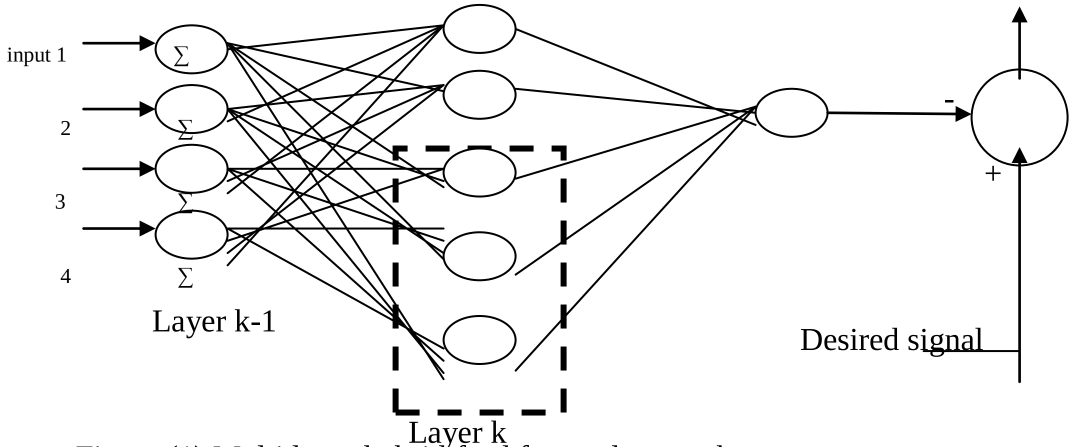


Figure (1) Multi layer hybrid feed forward network.

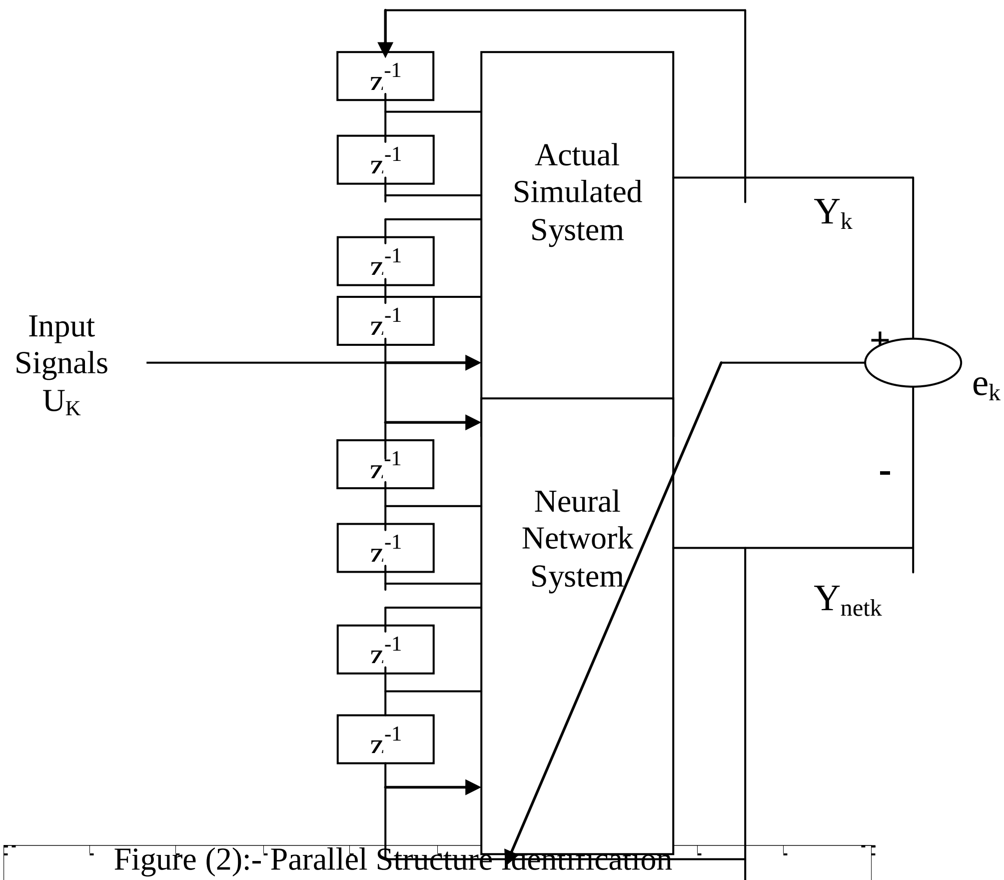
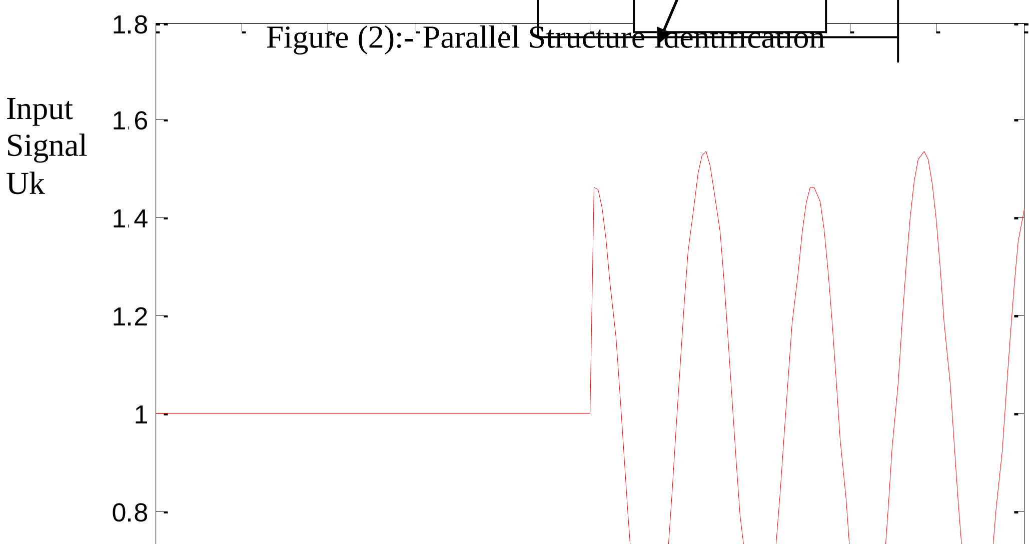


Figure (2):- Parallel Structure Identification



No. of Samples

Figure3:- Testing input signal U_k

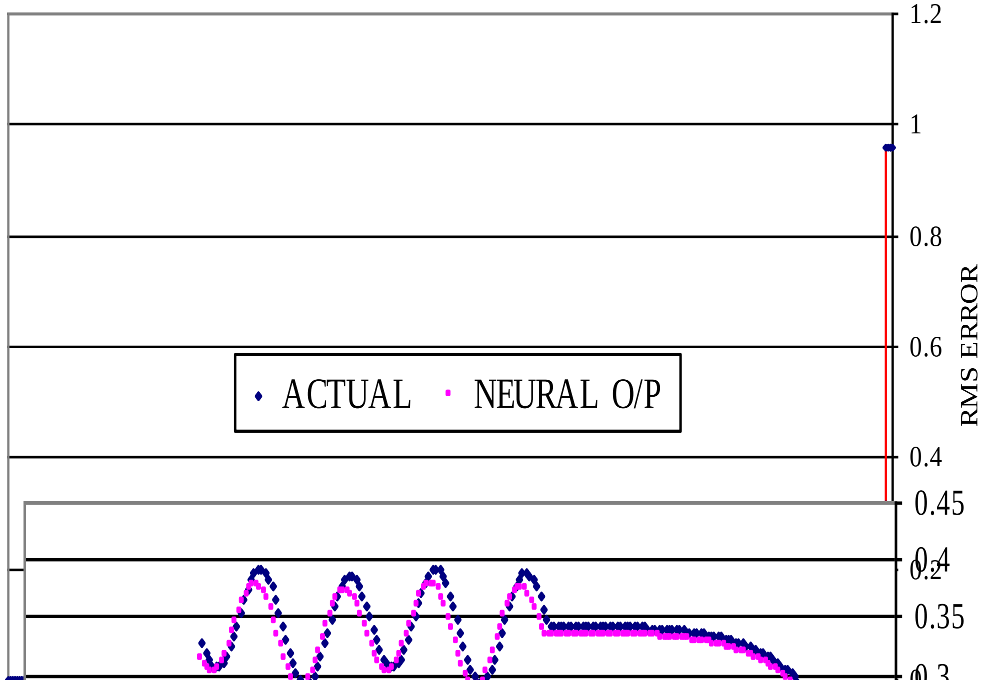


Figure 4:- RMS Error of the Plant1

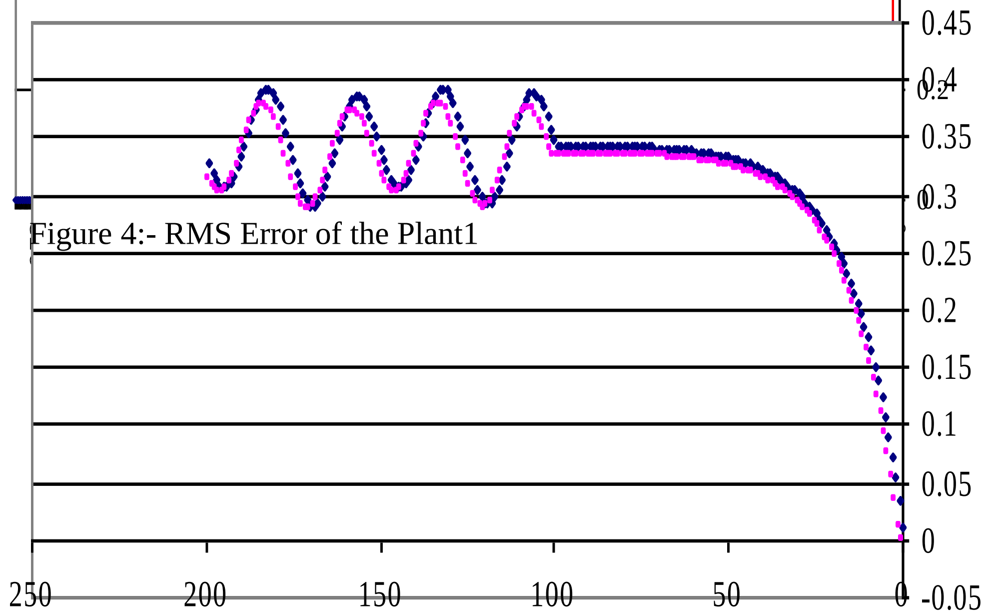
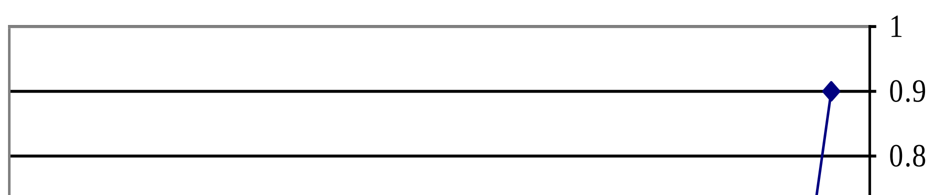


Figure5:-Plant1 and Neural Network Outputs To The Test Signal



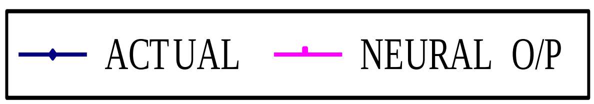


Figure6:- RMS Error of the Plant2

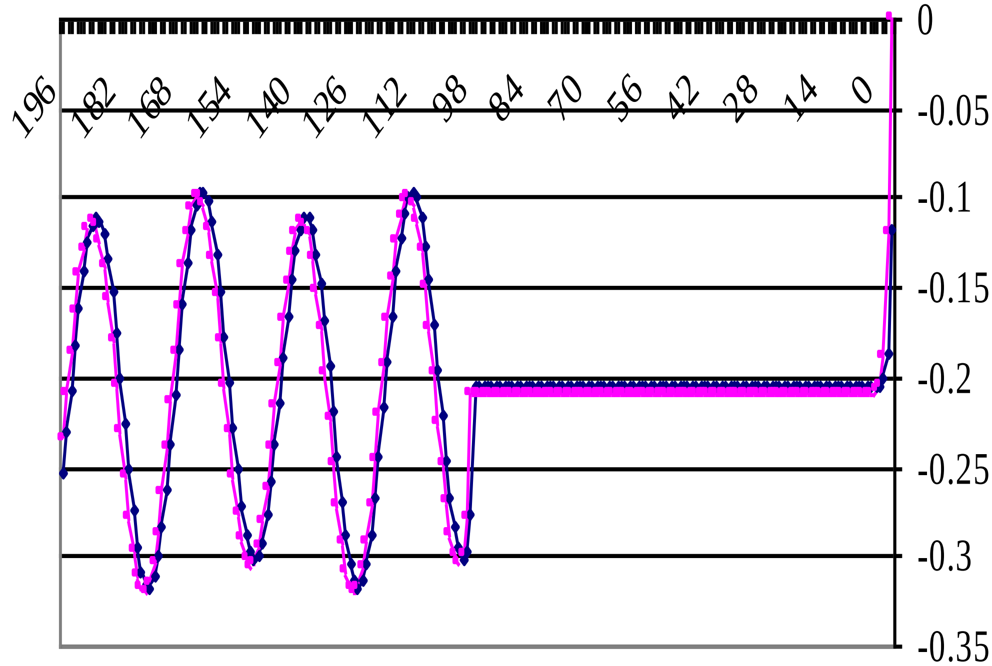
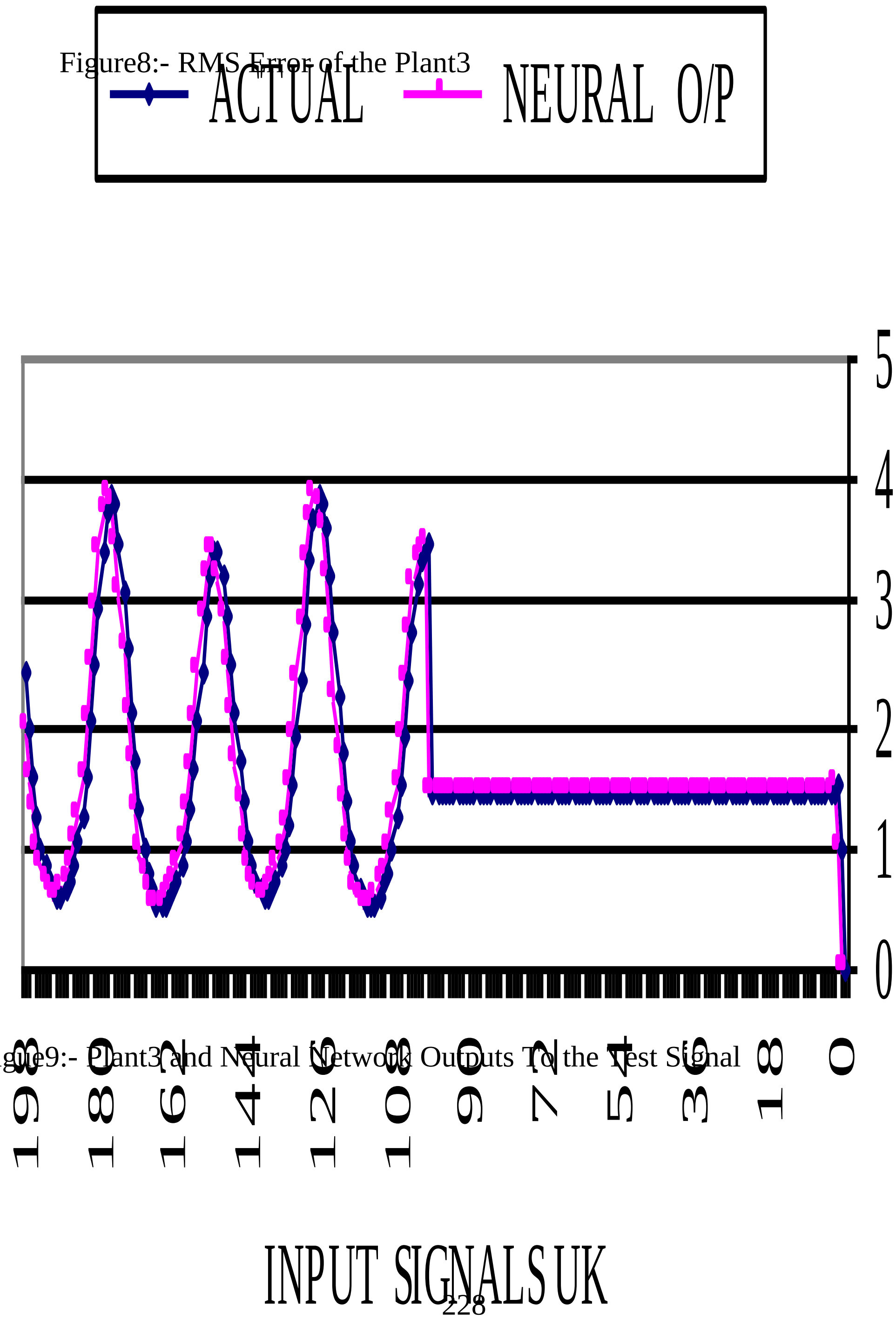


Figure7:-Plant2 and Neural Network Outputs To The Test Signal

INPUT SIGNALS UK

Figure8:- RMS Error of the Plant3



Figue9:- Plant3 and Neural Network Outputs To the Test Signal

References:-

1. DUC Truohg Pham and Liu Xing, “Neural Networks for Identification, Prediction and Control”, Springer-Verlag, 2nd, 1995.
2. Gary M. Scott and W. Harmon Ray, “Experiences with model-based controllers based on neural network process models”, Dept of Chemical Engineering, University of Wiscosin, Madison, USA, http//www.esf.edu,1993.
3. Gary M. Scott and W. Harmon Ray, “Neural Network Process Models Based on Linear Model Structure”, Dept of Chemical Engineering, University of Wiscosin, Madison, USA, http//www.esf.edu,1994.
4. Gary M. Scott and W. Harmon Ray, “Creating efficient nonlinear neural network process models that allow model interpretation”, Dept of Chemical Engineering, University of Wiscosin, Madison, USA, http//www.esf.edu,1993.
5. Fabio Codeca and Francesco Casella, “Neural Network Library in Modelica”, Politecnico di Milano, Italy, http// www.modelica.org,2006.
6. Xu Li, “Application of an identification method of BP neural networks in the speed control system of hydraulic elevator”, Zhejiang University of Technology, FIFTH INTERNATIONAL CONFERENCE ON FLUID POWER TRANSMISSION AND CONTROL(ICFP2001),3-5April,2001Hangzhou,China,http//www.fuid.power.net.
7. Paulo Gil, Jorge Henriquis, Antonio Dourado and H. Duarte-Ramos, “Extended neural model predictive control of non-linear systems”, http// www.cisuc.uc.pt,1998.
8. Jong-Sun Ko, Hyusik Kim and Seog-Hyun Yang,” Precision speed control of PMSM using neural network disturbance observer on forced nominal plant”, Dept of Electric Electronics and Computer Eng, Dankook University, http//www. ascc2004.ee.unimelb.edu.au,2004.
9. E.B. Kosmatopoulos, A.W. Smyth, S.F. Marsi and A.G.Ghassiako, “Adaptive neural identification of nonlinear structural systems”, http// www.ce.utexas.edu,2000.
10. Alma Y. Alanis, Edgar N. Sanchez and Alexander G. Loukianov, "Discrete-time recurrent neural induction motor control using kalman learning", http//www.fulton.asu.edu,2004.