THE EFFECT OF INPUT SEQUENCE TO NONLINEAR ARTIFICIAL NEURAL NETWORK PERFORMANCE IN MODELLING CSTR

Fariha Dwi Somanti , A.P.Dr.Suzana Yusup ,Haslinda Zabiri
Department of Chemical Engineering
University Teknologi PETRONAS
Bandar Sri Iskandar, 31750 Tronoh, Perak, Malaysia

Abstract- This paper considers the aspects of the system identification of nonlinear black box empirical models for chemical process dynamics. The core of this research is to study the effect of existing input sequence to nonlinear Artificial Neural Network applied in nonlinear dynamic system. To illustrate the practical utilization of the various types of input sequences used, NARXSP dynamic Neural Network model is applied to approximate the dynamics of a first-principles model of first order kinetic reaction in a simple Continued Stirred Tank Reactor.

Keywords: Input Sequence, Artificial Neural Network, NARXSP, Nonlinear System Identification, Dynamic System,CSTR

1. Introduction

Recently there are many development in input-output (I/O) modelling approach for dynamic modelling of nonlinear chemical processes, these include white box (fundamental/physical approach), grey box (semi empirical approach) or black box (empirical approach). However the two of the most important outstanding issues in significant development of I/O modelling are *the structure determination* and *input sequence design* (Chikkula,1997)

Artificial Neural Network (ANN) has proved to be able to solve complex tasks in a number of practical applications in chemical engineering such as fault detection, prediction of product quality, data rectification, modelling and control. In comparison to other empirical models are ANN can be used to solve highly nonlinear and complex sytem (Himmelblau,2000).

One of the model structure of ANN is NARX. NARX (Nonlinear Auto-Regressive with exogenous inputs) model structure has been proved as alternatives to nonlinear models in a number of chemical process applications such as distillation column (Ramesh, 2007), and CSTR (Doherty,1995; DeCicco,2000).

Selection of an appropriate input signal is considerably more challenging for nonlinear models (Leontarities , 1987). The input signal must have sufficient energy to excite the full range of nonlinear process dynamic. For linear systems identification, the Pseudo Random Binary Signal (PRBS) sequence have been commonly used. These signal cannot excite certain nonlinearities so that more input levels in the sequence are necessary (Doyle III,2002) . Gaussian White Noise Sequence is one of input sequence that is quite popular in providing adequate information about nonlinear process dynamic (Parker,2001). In the application of Neural Network, input sequence has important role for process excitation to generate input/output process data which constains sufficient information for a neural network model structure to identify the nonlinear process dynamic over the operating range (Doherty,1995)

This paper describes investigation into the effects of various input sequences with the aim of developing guidelines for nonlinear identification experiment design which will consequently improve the validity of the resulting neural network model . The investigation are carried out in a study of the development of multi-layered perceptron neural network model of first order kinetic reaction in a continuous stirred tank reactor (CSTR).

2. Simulated Continuous Stir Tank Reactor Model

For this research, simulation has done based on continuous stirred-tank reactor (CSTR) process model by Cervantes et al. Assuming that the process is irreversible, exothermic reaction, A→B, constant volume, and the reactor is cooled by a single coolant stream. The equation models are represented by equation (1) and (2).

$$C_A(t) = \frac{q(t)}{V}\left[C_{A0}(t) - C_A(t) - k_0\, C_A(t) exp\left[\frac{-E}{RT(t)}\right]\right] \quad (1)$$

$$T(t) = \frac{q(t)}{V}\left(T_0(t) - \frac{k_0 \Delta H}{\rho C_p} C_A(t) exp\left(\frac{-E}{RT(t)}\right)\right) + \frac{\rho_c C_{pc}}{\rho C_p V} q_c(t)\left[1 - exp\left(-\frac{hA}{q_c(t)\rho_c C_{pc}}\right)\right](T_{c0}(t) - T(t)) \qquad (2)$$

The objective of this CSTR is to control the output Temperature of the reactor as shown in Fig.1, by manipulating the coolant temperature T_c **(Cervantes et al, 2002)**.

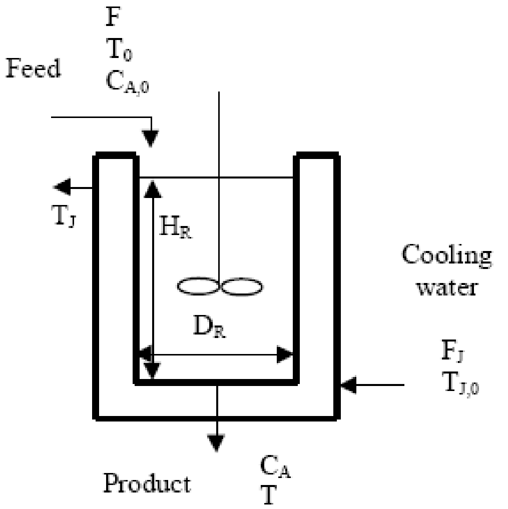

Fig.1. CSTR

3. Input Sequence

Several input sequences are used in nonlinear dynamic system that have sufficient energy to excite the full range of nonlinear process dynamic. A Random amplitude sequence is commonly used as the process excitation signal to generate open loop data for network training. This signal consist of a uniformly distributed random variable applied to the process input at each clock period and is more likely to exercise the process over the desired operating range than a binary signal (Pottman and Seborg, 1992). A Random amplitude signal is specified by its clock period, which should be a multiple of the sample time so that the process input is constant between consequtive samples, and by its amplitude range, which may be expressed as a percentage maximum deviation from a steady state value. (Doherty,1995). This sequence is shown by Fig.2. Gaussian White Noise (GWN) sequence, is completely random and uncorrelated with itself, which has the maximum possible input energy(Sony,2006). GWN sequence is shown in Fig.3. Multisine input sequence has been used in several case of highly nonlinear dynamic process and sufficient enough to excite the full range of the system.(Mittelmann, 2006; Rivera, 2003). This input sequence is represented by equation (3) and is shown in Fig.4.

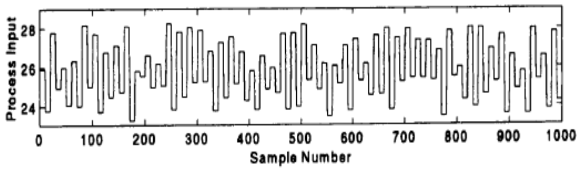

Fig.2 Random Amplitude Signal

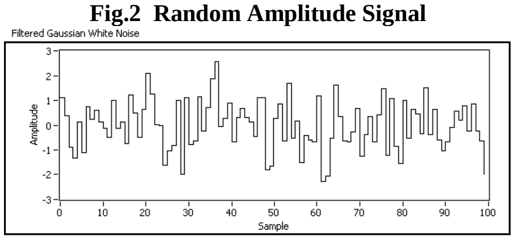

Fig.3 Gaussian White Noise Sequence input signal

$$\boxed{\begin{aligned} u_j(k) \;=\; & \sum_{i=1}^{m\delta} \hat{\delta}_{ji}\cos(\omega_i kT + \phi_{ji}^{\delta}) + \sum_{i=m\delta+1}^{m(\delta+n_s)} \alpha_{ji}\cos(\omega_i kT + \phi_{ji}) \\ & + \sum_{i=m(\delta+n_s)+1}^{m(\delta+n_s+n_a)} \hat{a}_{ji}\cos(\omega_i kT + \phi_{ji}^{a}), \;\; j = 1, \cdots, m \end{aligned}} \qquad (3)$$

Where T is sampling time, N_s is the sequence length, m is the number of channels, δ, n_s, n_a are the number of sinusoids per channel $(m(\delta+n_s+n_a)=N_s/2)$, ϕ^{δ}_{ji} , ϕ_{ji}, ϕ^{a}_{ji} , are the phase angles , a_{ji} represents the Fourier coefficient defined by the user, δ_{ji}, a_{ji} are the “snow effect” Fourier coefficients.

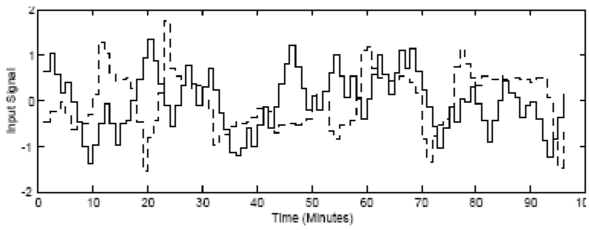

Fig.4 Multisine Input Sequence

In this paper, six types of sequences have been tested on the CSTR system to generate input-output data. For Sine input sequence , the signal resembles a sine curve (Figure 5). Theoretically, this is a simple method to obtain a wide range of the frequency responses by applying a wide range of input frequencies. However, in practice, this is not really applicable because the testing period can be extremely long during normal plant operation **(Tsai et al, 1986).** Generally, sinewave input sequence can be represented by equation (4).

$$U_{\sin}(t) = \begin{cases} 0 & t < 0 \\ A\sin\omega t & t \geq 0 \end{cases}$$

(4)

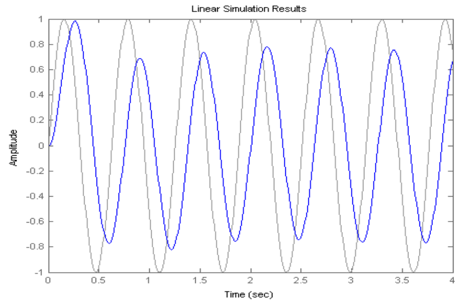

Fig 5. Sinewave input sequence

For Continuous switching–pace symmetric random sequence (CSRS) has a 4M +4 length deterministic input sequence used to estimate the bias, linear, and diagonal parameters. M is memory model of the system. This input sequences can be represented by equation (4) and (5).

$$u(k)=\begin{cases}\gamma_1 & k=0\\ 0 & 1\le k\le M\\ -\gamma_1 & k=M+1\\ 0 & M+2\le k\le 2M+1\\ \gamma_2 & k=2M+2\\ 0 & 2M+3\le k\le 3M+2\\ -\gamma_2 & k=3M+3\\ 0 & 3M+4\le k\le 4M+3\end{cases} \qquad (4)$$

$$u(k)=\begin{cases}\lambda_1 & k=1\\ \lambda_2 & k=2\\ 0 & 3\le k\le M+1\\ -\lambda_2 & k=M+2\\ -\lambda_1 & k=M+3\\ 0 & M+4\le k\le 2M+2\end{cases} \qquad (5)$$

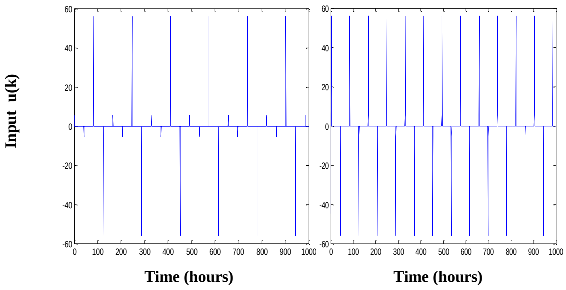

Fig.6 Input Sequence introduced by Sony et al and Parker et al

Input Sequence for Constant, linear, and diagonal parameters (2M+ 2-point input sequence) is four-level CSRS input sequence **(Parker et al, 2001**) represented by equation (6).

u(k) =

γ k = 0;

0 1 < k < M;

- γ k = M+1;

0 M+2 < k < 2M+1 (6)

where γ > 0, M=20

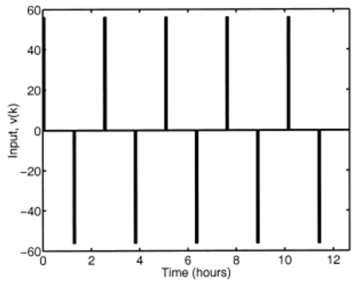

Fig.7 Input sequence introduced by Parker et al,2001

Signal of sum of sinusoids with different frequencies has introduced by **Baruch et al,2001** with specification of sequence has given in equation (7).

u(k) =

sin(πk/25) 0 < k < 26;

1.0 25 < k < 51;

-1.0 50 < k < 76;

0.3 sin(πk/25) + 0.1 sin(πk/32) +0.6 sin(πk/10), 75 < k < 101 (7)

Baruch et al also introduces sequence of pulses with random amplitude and width which is determined by equation (8).

u(k) =

sin(πk/25), 0 < k < 251;

1.0 250 < k < 501;

-1.0 500 < k < 751;

0.3 sin(πk/25) + 0.1 sin(πk/32) + 0.6 sin(πk/10), k < 1001 (8)

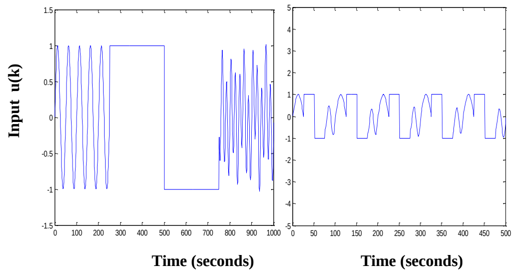

Fig.8 Input sequences introduced by Baruch et al,2001

Signal of random sequence has introduced by **Liu et al, 1997** is determined by equation (9).

u(k) =

sin(2πk/250), k ≤ 500;

0.8 sin(2πk/250) + 0.2 sin(2πk/25) k >500 (9)

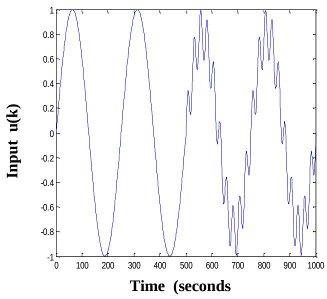

Fig.9 Input Sequence introduced by Liu et all,1997

4. NARX Model

NARX (Non-linear Auto-Regressive eXogenous) model also known as NARX Series-Pararel (NARX-SP) is recurrent neural network capable of modeling eficiently time series with long-term dependences is defined as

$$y(t)= f[y(t-1),...,y(t-n_y),u(t-k),...,u(t-k-nu+1)]+e(t) \qquad (10)$$

where *u(t)* and *y(t)* are the process input and output at time t,f[] is a non-linear function to be identified, *e(t)* is the equation error, k is the process deadtime (k>=l) and n_u, and n_y, are the number of delayed process inputs and outputs included in the model.

A neural network is configured in a NARX model structure by assigning the network input vector x = $[y(t-1) \; ..... \; y(t-n_y),u(t-k) \; ..... \; u(t-k-n_u+1)]$ and training the network to provide a one step ahead. prediction of the process output at time t, y(t). When the network has been trained, it can then be used to predict y(t+ 1) based on process data available at time t.

There are many applications for the NARX network. It can be used as a predictor, to predict the next value of the input signal. It can also be used for nonlinear filtering, in which the target output is a noise-free version of the input signal. The use of the NARX network is demonstrated in another important application that is modeling of nonlinear dynamic systems.

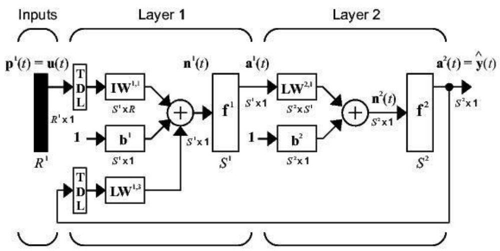

Figure 10 : NARXSP NN Architecture

Standard NARX architecture is as shown in Figure 7. The true output which is available during the training of the network is used instead of feeding back the estimated output. The advantage of this architecture is that the input to the feedforward network is more accurate. Besides, the resulting network has a purely feedforward architecture, and static back propagation can be used for training **(Zabiri and Mazuki, 2009)**.

Figure 8 describes the basic step for NARX model development using various input sequence.

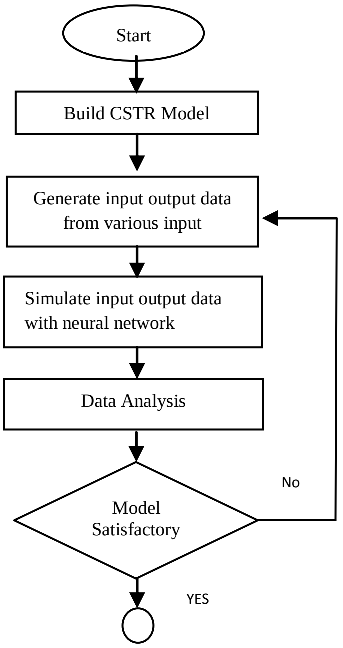

Fig.11 Basic Step for NARX model development

5. Result and Discussion

Simulated data generated from validated CSTR model using various types of input sequences are used to develop NARX model from the different input sequences. The seven of different input sequences are investigated.
Neural Network architecture is considered giving the best performance if it has low root mean squared error (RMSE), high correct directional change (CDC) and follows well the behavior of the signal.

The performance of Neural Network in predicting output Reactor Temperature under various type of input sequences is shown in Table 1.

Table 1. Comparison of NARXSP Prediction

Input Sequence	Neuron	RMSE	CDC	Level of Prediction	Transfer Function
Baruch et al	35	0.0161	80.203	well	PTL
Baruch et al	35	0.0174	91.371	merely	PTL
Liu et al	1	0.0030	80.203	well	LPP
Parker et al	50.	0.0136	89.848	slow	PPL
Soni et al	25.	0.0282	92.893	not well	TAT
Soni et al	6	0.0245	92.386	not well	PTP
Sinewave	35	0.0154	89.340	well	PLP

The best input sequences that can represent the best empirical model are Baruch 1, Sine and Liu inputs. From the input sequences that are tested in this study, Continuous Switching-pace Symmetric Random Sequence (CSRS) which are representated by Sony et al and Parker et al can not follow nonlinear behaviour of CSTR Dynamic System, since the CSRS has constant amplitude in longer sample time of k. Signal of sum sinusoids that are introduced by Baruch et al give good performance of NARXP model validation. The sum of sinusoids sequence that is using longer range of sample time of k shows better performance than the shorter ones. The signal of random sequence that is introduced by Liu has sinus sequence in the range of 500 sample time of k gives good performance in NARXSP model prediction , so does the sinewave sequence. From this study we can conclude that sinusoidal sequence apparently can excite full range of nonlinear dynamic system of CSTR with longer sample time of k.

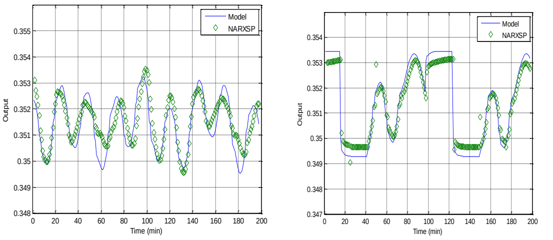

Fig.12 Neural Network Prediction with Baruch et al input sequence

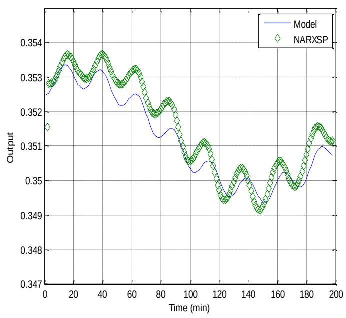

Fig.13 Neural Network Prediction introduced by Liu et al

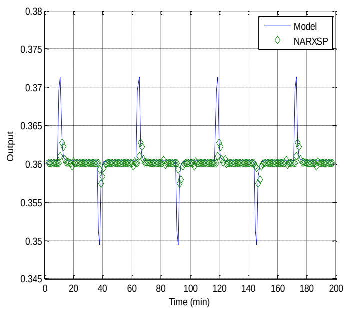

Fig.14 Neural Network Prediction introduced by Parker et al

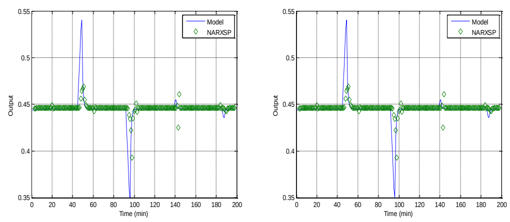

Fig.15 Neural Network Prediction introduced by Soni et al

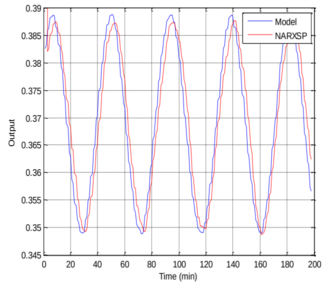

Fig.16 Neural Network Prediction with sinewave input sequence

.

In calculation, it can be concluded that NARXSP has been proved can capture for CSTR temperature profile satisfully and the type of sequences that introduced to the system have significant effect to NARXSP performance. Furthermore, it can be concluded that sinusoidal sequence is appropriate signal for nonlinear dynamic behaviour.

This paper only consider the significance of input sequence to excite full range of entire nonlinear dynamic system. However, the practicality of the input design is still an open issue. Future work is to design an input sequence feed is practical and “plant-friendly”.

6. References

[1] Chikkula.Y, Lee.J.H , Input Sequence Design for Parametric Identification of Nonlinear Systems, Proceedings of the American Control Conference(1997).

[2] DeCicco,J , Cinar.A , Empirical Modelling of Systems with Output Multiplicities by Multivariate Additive NARX Models, Ind.Eng.Chem,Res,39,6(2000).

[3] Doherty, S.K , Gomm, J.B and Williams.D ,Experiment Design Consideration for Non-Linear System Identification Using Neural Networks, Computers chem. Engng Vol.21, No.3(1997), 327-346 .

[4] Himmelblau, D.M. ,Applications of artificial neural networks in chemical engineering, Korean J.Chem.Res. 40 (2001) 373-392.

[5] Li,T. , Georgakis,C. ,Dynamic input signal design for the identification of constrained systems, Journal of Process Control 18 (2008) 332-346.

[6] Liu. G.P, Kadirkamanathan. V. Kadirkamanathan , S.A. Billings , On-line Identification of Nonlinear Systems Using Volterra Polynomial Basis Function Neural Networks, (1997).

[6] Mittelmann,H.D , Pendse.G, Optimal Input Design Signal in Data-Centric System Identification, Modern Mathematical Models, Methods and Algorithms for Real World Systems(2006).

[9] Parker, R.S, Heemstra.D, Doyle III,F.J, Pearson,R.K, Ogunnaike,B.A, The Identification of Nonlinear Models for process control using tailored :plant-friendly” input sequences, Journal of Process Control 11(2001) 237-250.

[10] Ramesh,K , Shukor, S.R A, Aziz.N , Development of Sigmoid Based NARX Model for a Distillation Column, ChemicalProduct and Process Modelling, Vol.3 (2008).

[11] Ramesh.K, Aziz.N , Shukor,S.R.A, Nonlinear Identification of Wavenet Based Hammerstein Model-Case Study on High Purity Distillation Column, Journal of Applied Sciences Research, 3(11), 1500-1508(2007).

[12]Pearson,R.K, Menold,P.H, Kraus.F.J, Set-Theoritic Input Sequence Design for Orthonormal Model Identification,

[13]Rivera,D.E, Lee.H, Braun,M.W, Mittelmann,H.D, "Plant-friendly" System Identification: A Chalengge for the process Industries, (2003).

[14]Valarmathi,K., Devaraj,D , Radhakrishan,T.K, Intelligent Techniques for System Identification and Controller Tuning in pH Process, Brazilian Journal of Chemical Engineering , 26(2009).

[15]Soni, A.S, Control –Relevant Sytem Identification using Nonlinear Volterra and Volterra-Laguerre Models,

[16]Yu,D.L, Gomm,J.B, Implementation of Neural Network predictive control to a multivariable chemical reactor, Control Engineering Practice 11(2003) 1315-1323

[17]Zabiri.H, Mazuki.N , A Black-Box Approach in Modeling Valve Stiction, International Journal of Natural Sciences and Engineering 2:1 (2009)

7. Acknowledgement

This research is supported by postgraduated office of Universiti Teknologi PETRONAS, Bandar Seri Iskandar, 31750 Tronoh, Perak, Malaysia

8. Notation

The Simulated temperature profile CSTR parameter values are given below.

Product Concentration	C_A	0.1 mol/l
Reactor Temperature	T	438.54 K
Coolant Flow Rate	Q_C	103.41 l/min
Process Flow Rate	Q	100 m^3/sec
Feed Concentration	C_{Af}	1mol/l
Feed Temperature	T_f	350 K
Inlet Coolant Temperature	T_{CO}	350 K
CSTR Volume	V	100 m^3
Heat Transer Term	H_A	7 X 10^5 cal/min K
Reaction rate constant	K_o	7.2 x 10^{10} l/min
Activation Energy Term	E/R	8750 K
Heat of Reaction	Δ H	5 x 10^4 J/mol
Liquid Densities	ρ, ρ_c	1000 kg/m^3
Specific heats	C_p, C_{pc}	0.239 J/kg K