Advanced Materials Research Vols. 785-786 (2013) pp 1430-1436
© (2013) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMR.785-786.1430

Identification of nonlinear systems using parallel Laguerre-NN model

H Zabiri[1, a], M Ramasamy[1,b], T D Lemma[1,c] and A Maulud[1,d]

[1]Chemical Engineering Dept, Universiti Teknologi PETRONAS, Bandar Seri Iskandar, 31750 Tronoh, Perak, MALAYSIA

[a]haslindazabiri@petronas.com.my, [b]marappagounder@petronas.com.my, [c]lemma_dendena@petronas.com.my, [d]halims@petronas.com.my

Keywords: Nonlinear system identification, orthonormal basis filters, neural networks

Abstract. In this paper, a nonlinear system identification framework using parallel linear-plus-neural networks model is developed. The framework is established by combining a linear Laguerre filter model and a nonlinear neural networks (NN) model in a parallel structure. The main advantage of the proposed parallel model is that by having a linear model as the backbone of the overall structure, reasonable models will always be obtained. In addition, such structure provides great potential for further study on extrapolation benefits and control. Similar performance of proposed method with other conventional nonlinear models has been observed and reported, indicating the effectiveness of the proposed model in identifying nonlinear systems.

Introduction

Nonlinear system identification has been the subject of intense research in recent years [1]. In general, the prediction of the output at time *t* of a dynamical system can be represented as [1]:

$$\hat{y}(t|\theta) = g(Z^{t-1}, \theta) \tag{1}$$

where $Z^{t-1} = \{y(t-1), u(t-1), y(t-2), u(t-2), \ldots\}$ is the vector of all or some previously measured inputs and outputs, and in cases where the system is not fully known, the prediction will be parameterized by the parameters θ. A nonlinear dynamic model is when $\hat{y}(t|\theta)$ is nonlinear in Z and nonlinear/linear in θ.

In the control community, one of the important issues for dynamic systems is the selection of the useful parameterizations of $\hat{y}(t|\theta)$ for nonlinear models [1]. Due to the fact that in nonlinear process modeling there exists no universal procedure for representing a nonlinear process, a vast amount of modeling approaches have been investigated [2, 3]. In essence, there are two standard approaches for building mathematical models: the traditional fundamental modeling (white box) and the purely empirical modeling (black box) [1, 4]. Grey box or hybrid model is the term used when these two approaches are combined.

Over the years numerous nonlinear empirical models have been developed, and they can generally be classified into two major categories: (1) single structure-based empirical models, and (2) linear-and-nonlinear-based empirical models. Many of the widely known modeling techniques fall under the single structure-based models category, which include Volterra models, artificial neural networks, fuzzy-logic based models, and Nonlinear Auto Regressive with eXternal input (NARX) models. Some combinations of them like neuro-fuzzy models, support vector machine and kernel methods of modeling, and wavelet decomposition based models are also reported. On the other hand, linear-and-nonlinear-based empirical models may be classified into series and parallel forms. The most popular type of the series form is the block-oriented (BO) models, for example, Hammerstein and Wiener models [5-12].

In nonlinear system identification using black box models such as neural networks, one possible approach is to use a parallel combination of linear-plus-neural networks models. The residuals from a linear model may be used to develop a neural networks model to pick up the nonlinearities [13]. This parallel structure ensures that reasonable models are always obtained and the overall model performs at least as good as or better than the linear model [14]. An interesting approach is the two-point gain-scheduling method using a static neural network model and a quadratic difference equation [15]. Another approach is the use of integrated linear partial least-squares (PLS) and nonlinear static feed-forward neural network in parallel in a structure known as extended Wiener model [16]. However, methods used involved complex procedures. An integrated linear state space model with neural networks model structure has also been reported for the identification and control of a one-degree-of-freedom vibration system [17], however, the model is only applicable for mild nonlinear systems. In this paper, a simpler approach is proposed with the development of a nonlinear model is proposed using the integration of parallel linear Laguerre filters model and neural networks (NN) model.

Parallel Laguerre-NN model

Consider a general nonlinear output error (NOE) model structure expressed as

$$y(k) = f(u(k-1),...,u(k-m),\hat{y}(k-1),\ldots,\hat{y}(k-m)) + e(k) \quad (2)$$

where $e(k)$ refers to the system white noise. A general linear model structure, on the other hand, may be represented as

$$y(k) = G(q)u(k) + e(k) \quad (3)$$

Without loss of generality (2) and (3) can be combined to get

$$y(k) = G(q)u(k) + f(u(k-1),...,u(k-m),\hat{y}_r(k-1),\ldots,\hat{y}_r(k-m)) + e(k) \quad (4)$$

where $\hat{y}_r$ refers to the predicted residuals of the linear model, *i.e.* $\hat{y}_r = y_{measured} - \hat{y}_{linear}$. Equation (4) represents a parallel structure in which a linear model is combined with a non-linear model represented by $f(\cdot)$. For a single-input single-output system with a Multi-Layer Perceptron (MLP) neural network with one hidden layer in parallel with a linear Laguerre filter model, the one-step ahead prediction model from (4) becomes

$$\hat{y}(k+1) = \left(\sum_{j=1}^{N} c_j L_j(q)\right)u(k) + \beta\left[b^2 + \sum_{i=1}^{K} w_i^2 \varphi(b_i^1 + w_{i,1}^1 x(k))\right] \quad (5)$$

where the nonlinear neural network function approximation is trained with regression vectors consisting of previous plant inputs and previous residuals of the linear model, $x(k) = [u(k-1),\ldots,u(k-m), y_r(k-1),\ldots,y_r(k-m)]$. Also $\varphi, \beta : R \to R$ are the nonlinear activation functions (in this research work, hyperbolic tangent activation function is used for the hidden layer), b are the biases, K is the number of hidden neurons, and the weights of the network are denoted by $w_{i,j}^1, i = 1,\ldots,K$ (with i^{th} neuron and j^{th} input, in this case $j = 1$) for the first layer, and $w_i^2, i = 1,\ldots,K$ for the second layer.

Nonlinear identification algorithm. The sequential identification structure proposed for the parallel Laguerre-NN models is illustrated in Figure 1. The linear Laguerre model is identified first, and the nonlinear NN model is then trained with the predicted residuals. The *pseudo*-independent nature of this parallel structure allows both the models to capture the essential characteristics of the underlying process separately and hence more accurately.

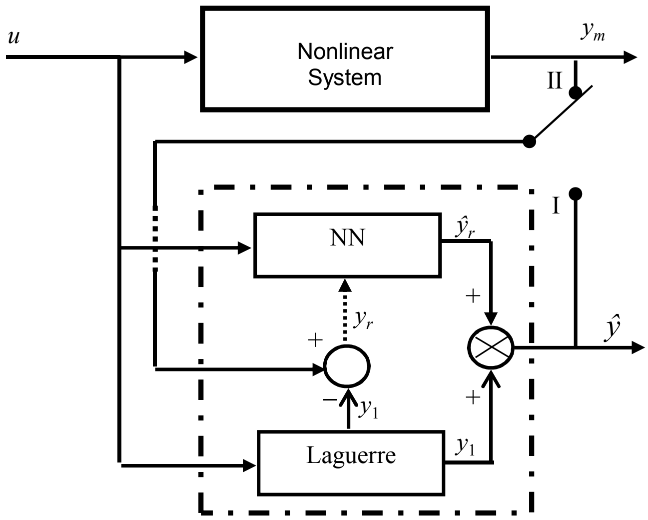


Fig. 1. The proposed sequential identification of residuals-based parallel Laguerre-NN models (I: simulation configuration, II: prediction configuration)

Given a set of nonlinear data to be identified $[u(k), y_m(k)]$, the simple algorithm to identify the Laguerre-NN model can be described as follows:

Algorithm 1

1. Develop a parsimonious Laguerre model using methods described by [18] to get y_1.
2. Calculate the predicted residuals using $\hat{y}_r = y_m - y_1$.
3. Develop the MLP network model using standard algorithm with $x(k) = [u(k-1),\ldots,u(k-m),\hat{y}_r(k-1),\ldots,\hat{y}_r(k-m)]$ as inputs and $\hat{y}_r(k)$ as outputs of the NN model.

Results and Discussions

Case study description. The system considered is a simple single-input single-output (SISO) nonlinear plant frequently referred to in literature for nonlinear systems analysis and control [19], with process input, u, and process output, y. The model is given by:

$$y(k) = \frac{2.5y(k-1)y(k-2)}{1+y^2(k-1)+y^2(k-2)} + 0.3\cos(0.5(y(k-1)+y(k-2))) + 1.2u(k-1) \qquad (6)$$

A slight modification is introduced to induce a higher nonlinear behaviour, where cosine function is imposed on the process input term as shown in (7):

$$y(k) = \frac{2.5y(k-1)y(k-2)}{1+y^2(k-1)+y^2(k-2)} + 0.3\cos(0.5(y(k-1)+y(k-2))) + \underline{\underline{\cos(1.2u(k-1))}} \qquad (7)$$

This nonlinear plant has a pre-specified sampling time which is 1 time unit. For this case study, a sinusoidal input signal is used to generate the identification data. The sinusoidal signal is generated using the 'sine' block in SIMULINK, with amplitude of 1.0 and frequency of 0.025 sampled at 1.0 time unit for 500 data points.

Identification. The input-output data for the identification of this case study is given in Figure 3.

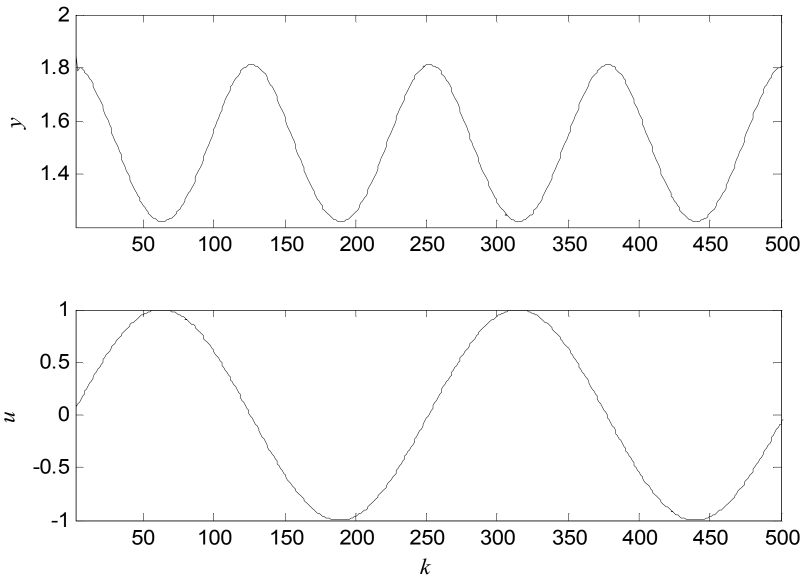
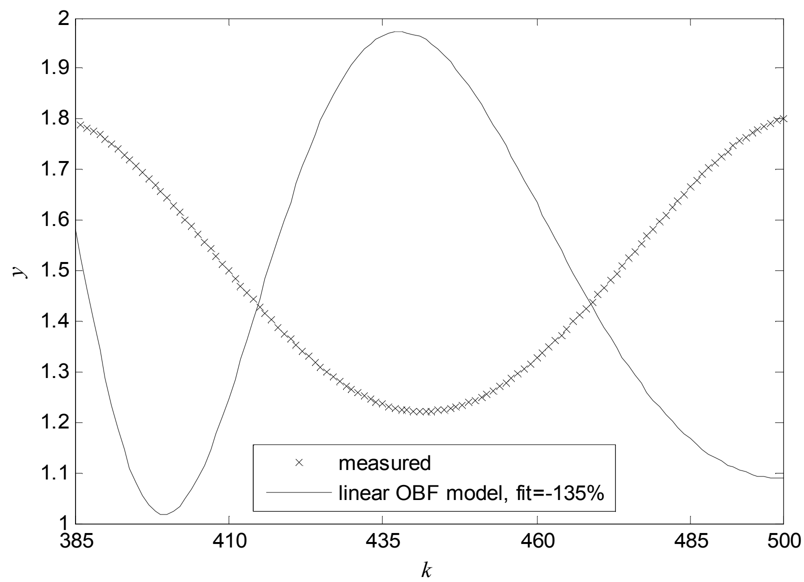


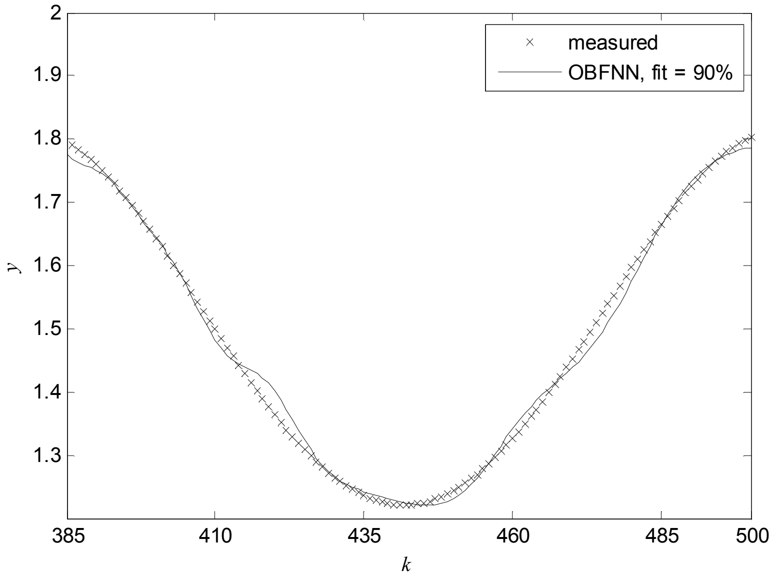
Fig.3. Input-output data set

The Laguerre model for the Laguerre-NN is developed with 6 Laguerre filters. The estimated pole and Laguerre parameters are

$\hat{p} = 0.9656$
$c_{OBF\text{-}NN} = [0.0641\ \text{-}2.5678\ 0.6077\ 0.7902\ \text{-}0.5826\ 1.1567]$



(a) Measured and estimated output for linear Laguerre model (validation data)



(b) Measured and predicted output for the overall Laguerre-NN model (validation data)
Fig.4. Laguerre-NN model performance on validation data set

Figure 4(a) shows the corresponding linear Laguerre model performance on the validation data set for this plant. It is clear that the plant considered is highly nonlinear, as the percentage of fit of the linear OBF model is well below 0%. Once the linear Laguerre model has been developed and the residuals have been calculated, $\hat{y}_r = y_m - y_1$, the residuals MLP networks are then trained. Hyperbolic tangent activation function (represented as 'tansig') is selected for the hidden layer, and the output layer is set at linear. The resulting residuals network configurations may be represented as follows:

res_{NN}: [4-4-1 neurons with tansig-tansig-linear transfer functions]

The overall Laguerre-NN model performance is shown in Figure 4(b). Comparing Figures 4(a) and (b), significant improvement can be seen in the overall Laguerre-NN identified model. The Laguerre-NN model is able to track the true output behavior as shown in Figure 4(b). For comparison analysis, models based on conventional MLP (NN) and series Wiener-MLP are used. The identified conventional NN model has 4 neurons in its hidden layer with the following configuration:

$NN_{(plant1)}$: [4-4-1 neurons with tansig-tansig-linear transfer functions]

The resulting linear Laguerre subsystem for the Wiener-MLP model with 6 Laguerre filters has the following estimated pole and Laguerre parameters

$\hat{p} = 0.9733$

$c_{Wiener\text{-}MLP(plant1)}$ = [0.5618 3.6426 -6.2570 3.2004 -2.7261]

The corresponding nonlinear MLP NN subsystem has 22 hidden neurons in its hidden layer with the following optimal configuration:

$seriesNN_{(plant1)}$: [4-22-1 neurons with tansig-tansig-linear transfer functions]

The performance comparisons are done on the validation data set. Figure 5 provides the output comparisons for the Laguerre-NN, conventional NN, and the series Wiener-MLP models. It is observed that the performances of all three models are fairly similar with each other, indicating the capability of the proposed Laguerre-NN model in achieving almost comparable accuracy in identifying the nonlinear system as the conventional NN and Wiener-MLP models. In addition, this algorithm requires only one type of input design sequence for the identification of both linear OBF and the MLP network. Identification of each model also requires the usage of well-established methods without suffering from any other added complexity. Besides these advantages, the parallel structure also has great potential for extrapolation since in the worst scenario linear extrapolation behaviour can be expected.

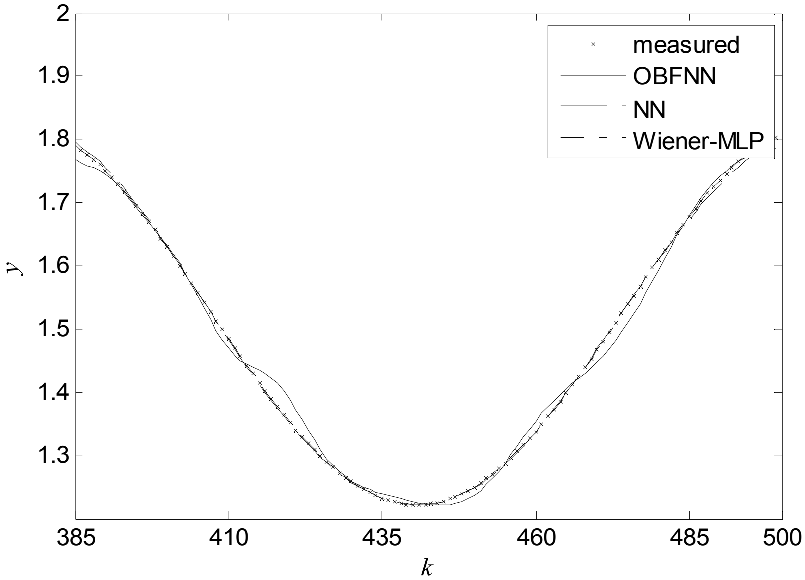


Fig. 5. Output comparison (validation data)

Conclusions

In this paper, parallel Laguerre-NN model has been developed and tested for the identification of a nonlinear plant system. The results indicate the capability of the proposed Laguerre-NN model in identifying the nonlinear system and providing comparable performance with other conventional methods. This is promising as the parallel structure has great potential for further study on extrapolation benefits and control.

Acknowledgements

This work was financially supported by Universiti Teknologi PETRONAS.

References

[1] Ljung, L., *Perspectives on system identification.* Annual Reviews in Control, 2010. **34**(1): p. 1-12.

[2] Ou, J. and Rhinehart, R.R., *Grouped-neural network modeling for model predictive control.* ISA Transactions, 2002. **41**(2): p. 195-202.

[3] Ou, J. and Rhinehart, R.R., *Grouped neural network model-predictive control.* Control Engineering Practice, 2003. **11**(7): p. 723-732.

[4] Abdelazim, T. and Malik, O.P., *Identification of nonlinear systems by Takagi–Sugeno fuzzy logic grey box modeling for real-time control.* Control Engineering Practice, 2005. **13**(12): p. 1489-1498.

[5] Aadaleesan, P., Miglan, N., Sharma, R., and Saha, P., *Nonlinear system identification using Wiener type Laguerre–Wavelet network model.* Chemical Engineering Science, 2008. **63**(15): p. 3932-3941.
[6] Arefi, M.M., Montazeri, A., Poshtan, J., and Jahed-Motlagh, M.R., *Wiener-neural identification and predictive control of a more realistic plug-flow tubular reactor.* Chemical Engineering Journal, 2008. **138**(1–3): p. 274-282.
[7] Arto, V., Hannu, P., and Halme, A., *Modeling of chromatographic separation process with Wiener-MLP representation.* Journal of Process Control, 2001. **11**(5): p. 443-458.
[8] Hsu, Y.-L. and Wang, J.-S., *A Wiener-type recurrent neural network and its control strategy for nonlinear dynamic applications.* Journal of Process Control, 2009. **19**(6): p. 942-953.
[9] Kalafatis, A., Arifin, N., Wang, L., and Cluett, W.R., *A new approach to the identification of pH processes based on the Wiener model.* Chemical Engineering Science, 1995. **50**(23): p. 3693-3701.
[10] Lawryńczuk, M., *Computationally efficient nonlinear predictive control based on neural Wiener models.* Neurocomputing, 2010. **74**(1–3): p. 401-417.
[11] Mahmoodi, S., Poshtan, J., Jahed-Motlagh, M.R., and Montazeri, A., *Nonlinear model predictive control of a pH neutralization process based on Wiener–Laguerre model.* Chemical Engineering Journal, 2009. **146**(3): p. 328-337.
[12] Shafiee, G., Arefi, M.M., Jahed-Motlagh, M.R., and Jalali, A.A., *Nonlinear predictive control of a polymerization reactor based on piecewise linear Wiener model.* Chemical Engineering Journal, 2008. **143**(1–3): p. 282-292.
[13] Sjöberg, J., Zhang, Q., Ljung, L., Benveniste, A., Delyon, B., Glorennec, P.-Y., Hjalmarsson, H., and Juditsky, A., *Nonlinear black-box modeling in system identification: a unified overview.* Automatica, 1995. **31**(12): p. 1691-1724.
[14] Nelles, O., *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*. 2001, Berlin, Germany: Springer.
[15] Piche, S., Sayyar-Rodsari, B., Johnson, D., Gerules, M., *Nonlinear Model Predictive Control using Neural Networks*, in *Proc. IEE Contr. Syst. Mag.* . 2000. p. 53-62.
[16] Zhao, H., Guiver, J., Neelakantan, R., and Biegler, L.T., *A nonlinear industrial model predictive controller using integrated PLS and neural net state-space model.* Control Engineering Practice, 2001. **9**(2): p. 125-133.
[17] Yasui, T., Moran, A., Hayase, M. . *Integration of linear systems and neural networks for identification and control of nonlinear systems*. in *Proc. 35th SICE Annual Conference (SICE'96)*. 1996. Tottori.
[18] Tufa, L.D., *Control Relevant System Identification Using Orthonormal Basis Filter Models*, in *Chemical Engineering*. 2009, UNIVERSITI TEKNOLOGI PETRONAS: Tronoh.
[19] Lazar, M., Pastravanu, O. , *A neural predictive controller for nonlinear systems.* Mathematics and Computers in Simulation, 2002. **60**: p. 315-324.