Journal of Applied Sciences

ISSN 1812-5654

science alert

ANSInet
an open access publisher
http://ansinet.com

Journal of Applied Sciences 14 (22): 3063-3069, 2014
ISSN 1812-5654 / DOI: 10.3923/jas.2014.3063.3069

Multi-Step Ahead Prediction Analysis for MPC-Relevant Models

H. Zabiri, M. Ramasamy, L.D. Tufa and A. Maulud
Department of Chemical Engineering, Universiti Teknologi PETRONAS, Bandar Seri Iskandar, Tronoh, 31750, Perak, Malaysia

Abstract: Model Predictive Control (MPC) is one of the most successful controllers in industries and widely applied in petroleum refining and petrochemical processes. Its inherent model-based strategy, however, renders it sensitive to changes that occur when the plants operate outside the boundaries of its original operating conditions. In this study, a nonlinear empirical model based on parallel orthonormal basis function-neural network structure, which has been shown to be able to extend the applicable regions of the model, is evaluated for its multi-step ahead prediction capability and compared to the conventional neural network models with different scaling procedures. It has been shown that the nonlinear model exhibited sufficient multi-step ahead prediction capability that renders it a promising candidate for MPC applications that can potentially improve the closed-loop control performance in extended regions and this is important in retaining the positive benefits of MPC in industries.

Key words: Petroleum refining, MPC, advanced process control, linear and non-linear model

INTRODUCTION

Industries nowadays are becoming large and complex and highly energy-intensive. Petroleum refining processes, in particular, consume substantial energy and are considered as the most energy-intensive manufacturing industry in the US. One of the technologies that make existing assets more energy efficient is Advanced Process Control (APC) (Wakasugi and Sueyoshi, 2010). However, Model Predictive Control (MPC), which is the heart of APC system, are generally sensitive to any changes in the process conditions on which the original implementation was based. Any corresponding adjustments need to be made to the models requires re-testing and re-modeling which are very expensive from maintenance point of view. Hence, MPC-relevant nonlinear empirical models with extended applicable regions are highly desirable so that subsequent reduction in maintenance cost may be achieved.

Linear-and-nonlinear-based empirical models seem to be an interesting alternative for extending the applicable regions of the model (i.e., extrapolation enhancement) in terms of the wider range of possible combinations that can be explored. In particular, the parallel integration of linear and nonlinear models holds much potential as the usage of residuals of the linear model ensures that a reasonable model is always obtained, the overall nonlinear model will always be as good as or better than the linear model used and the opportunity of the underlying nonlinear characteristics to be captured by the residuals (Ardalani-Farsa and Zolfaghari, 2010; Nelles, 2001; Sjoberg *et al*., 1995).

An interesting 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 (Zhao *et al*., 2001). The methods used, however, involved complex procedures. Another approach is the two-point gain-scheduling method using a static neural network model and a quadratic difference equation (Piche *et al*., 2000). Though excellent closed-loop performances have been shown in both (Zhao *et al*., 2001; Piche *et al*., 2000) under various set-point and load changes, no extrapolation analysis is discussed. The objective function used in the nonlinear optimization problem of Piche *et al*. (2000) also contains higher number of terms in comparison to the standard two terms (desired outputs and input move suppressions) normally encountered in nonlinear MPC application. 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 (Yasui *et al*., 1996) however, no analysis is considered in the extended regions of the developed model.

A recently developed parallel Orthonormal Basis Function (OBF), Neural Networks (NN) model have shown significantly improvement in extending the applicable

Corresponding Author: H. Zabiri, Department of Chemical Engineering, Universiti Teknologi PETRONAS, Bandar Seri Iskandar, Tronoh, 31750, Perak, Malaysia Tel: +6053687625

regions of the model under open-loop conditions in comparison to other conventional models (Zabiri *et al*., 2013). Despite the promising benefits, however, the results reported are based on open-loop simulation studies only. For any model to be useful in the nonlinear model-based predictive controller framework, an important facet is the model capability to perform satisfactory multi-step ahead predictions. Hence, this study aims to evaluate and study the multi-step ahead prediction analysis of the parallel OBF-NN model and compared the results with conventional neural networks models scaled using standard scaling methods and spread encoding techniques.

ALGORITHMS AND METHODS

Parallel OBF-NN multi-step ahead predictor: The methods of calculating multi-step ahead prediction using OBF-NN model is based on the recursive prediction strategy. Only one model is utilized for making the prediction in the recursive prediction strategy (Cheng *et al*., 2006; Girard *et al*., 2002). A sequential algorithm based on the knowledge of current values of u and y together with the OBF-NN system model gives the h-step-ahead prediction. For a single input single output system with a multi-layer perceptron neural network with one hidden layer in parallel with a linear OBF model, overall equation is given as:

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

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-1) = [u (k-1), ..., u (k-m), $\hat{y}_r$ (k-1), ... , $\hat{y}_r$ (k-m)]. Also φ, γ: R→R are the nonlinear activation functions (e.g., hyperbolic tangent etc.), 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 for the second layer:

$$w_i^2,\ i=1,\ \ldots,K$$

N is the number of orthonormal basis filters, c are the optimal OBF model parameters, L_j (q) are the orthonormal basis filters and q is the forward shift operator.

From Eq. 1, the OBF-NN model prediction at the k+1 time instant can be derived as:

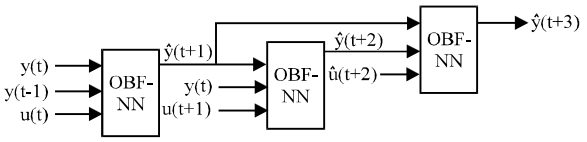

Fig. 1: OBF-NN -based multi-step ahead predictor employing recursive prediction method

$$\hat{y}(k+1)_{OBFNN}=\hat{y}_{OBF}(k+1)+\beta[b^2+f(x(k))] \quad (2)$$

Extending the prediction one-step further, ŷ(k+2) can be obtained and generally, the h-step ahead predictor can be derived as follows:

$$\hat{y}(k+h)_{OBFNN}=\hat{y}_{OBF}(k+h)+\beta[b^2+f(x(k+h-1))] \quad (3)$$

Where:

$$x(k+h-1)=[u(k+h-1),\ \ldots,\ u(k+h-m),\ \ldots\ \hat{y}_r(k+h-1),\ \ldots,\ \hat{y}_r(k+h-m)] \quad (4)$$

In Eq. 4, the output sequence for h-step ahead prediction for the linear OBF model y_{OBF} (k+h) is a function of the input sequence up to instant u(k+h-1) (Tufa and Ramasamy, 2009). The multi-step ahead predictors for the residuals network in Eq. 4 also adopted the recursive prediction strategy (Gomm *et al*., 1996).

Figure 1 illustrates an example of OBF-NN that acts as a multi-step predictor under the assumption of m_y = 2, m_u = 1, d = 1 and prediction horizon P = 3.

Neural networks based multi-step ahead predictor: The output of the nonlinear NN model can be given as follows:

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

where, the regression vectors consisting of previous plant inputs and previous process outputs. From Eq. 5, the NN model output at the k+h time instant can be derived as:

$$\hat{y}(k+h)_{NN}=\beta[b^2+f(x(k+h-1))] \quad (6)$$

Where:

$$x(k+h-1)=[\hat{y}(k+h-1),\ \ldots,\ \hat{y}(k+h-m),\ \ldots\ u(k+h-1),\ \ldots,\ u(k+h-m)] \quad (7)$$

Similar to the residuals network in previous section, the multi-step ahead prediction for the conventional NN is done via recursive prediction.

The multi-step ahead prediction capability of the proposed OBF-NN model is to be tested before the model is deemed suitable to be applied in the NMPC framework. With regards to this objective, the comparison is to be made with the conventional NN network developed using two different scaling techniques: (1) Standard scaling method which has been widely used in various applications and (2) Spread encoding technique proposed by Gomm *et al*. (1996). The second method is introduced as a possible measure of improvement (if any) for the long range prediction capability of conventional MLP neural network model.

In the standard scaling method, the following equation is used to normalize the input-output data into the range of [-1,1] using:

$$r_{norm} = 2\left(\frac{r - r_{min}}{r_{max} - r_{min}}\right) - 1 \quad (8)$$

The spread encoding technique is described in detail by Gomm *et al*. (1996). Assume that the original data range is given by $r \in [r_{min}, r_{max}]$. Also, N_t is the total number of nodes used for the coding and decoding steps, N_0 is the number of nodes on either side of the variable range and $\delta = 1$. The Spread-encoding method to code and decode a value r involves the following steps:

Encoding steps:

Step 1: Scale r to the normalized range by:

$$x_{SE} = \frac{N_t - 2N_0}{r_{max} - r_{min}}(r - r_{min}) \quad (9)$$

Step 2: Code the data to the N_t network nodes by:

$$\Psi_i(x_{SE}) = \Phi(a_i + \frac{1}{2} - x_{SE}) - \Phi(a_i - \frac{1}{2} - x_{SE}),\ i = 1, ..., N_t \quad (10)$$

where, $a_i = i - N_0 - c_c$ and c_c is called the offset term that shifts the position of the range limits on the nodes and is set to a value of 0.5 as it is found to be sufficient for accurate coding and decoding process. In addition, the total number of nodes and the number of nodes on either side of the variable range are set to $N_t = 6$ and $N_0 = 2$, respectively. Also, $\Phi(a)$ is the sigmoidal function centred at x_{SE}:

$$\Phi(a - x_{SE}) = \frac{1}{1 + e^{-\gamma_{SE}(a - x_{SE})}} \quad (11)$$

where, γ_{SE} values control the width of the node excitations. It is found that for the work related in this study, $\gamma_{SE} = 2$ provides sufficient accuracy for the coding and decoding process:

Step 3: Scale the excitation of each node to the range [0.1, 0.9] using a fixed linear relationship:

$$\bar{\Psi}_i(x_{SE}) = \left(\frac{\Psi_i(x) - \Psi_{min}}{\Psi_{max} - \Psi_{min}}\right)(0.8) + 0.1 \quad (12)$$

Decoding steps:

Step 1: Apply the inverse of the scaling relationship used in Step 3 of the coding procedure to descale the node excitations from the range [0.1, 0.9] back to their original range:

$$\Psi_i(x_{SE}) = \left(\frac{\bar{\Psi}_i(x_{SE}) - 0.1}{0.8}\right)(\Psi_{max} - \Psi_{min}) + \Psi_{min} \quad (13)$$

Step 2: Decode the network output back to the normalized range using:

$$x_{SE} = \frac{\sum_{i=1}^{N_t} a_i \Psi_i(x_{SE})}{\sum_{i=1}^{N_t} \Psi_i(x_{SE})} \quad (14)$$

Step 3: Apply the inverse of the scaling relationship used in Step 1 of the coding procedure to determine the final decoded value r:

$$r = \left(\frac{r_{max} - r_{min}}{N_t - 2N_0}\right)x_{SE} + r_{min} \quad (15)$$

RESULTS AND DISCUSSIONS

Case study description: The van de Vusse reactor (Vojtesek *et al*., 2004) is shown in Fig. 2. A reactant A is to be converted to the desired product B, but the product B degrades to product C. In addition to these consecutive reactions, a high-order parallel reaction occurs by which the reactant A is converted to by-product D:

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C$$

$$2A \xrightarrow{k_3} D$$

The mathematical model of this reactor is described by the following set of Ordinary Differential Equations (ODE):

$$\frac{dc_A}{dt} = \frac{q_r}{V_r}(c_{A0} - c_A) - k_1 c_A - k_3 c_A^2 \quad (16)$$

$$\frac{dc_B}{dt} = -\frac{q_r}{V_r}c_B + k_1 c_A - k_2 c_B \quad (17)$$

$$\frac{dT_r}{dt} = \frac{q_r}{V_r}(T_{r0} - T_r) - \frac{\Delta H_r}{\rho_r c_{pr}} + \frac{A_r U}{V_r \rho_r c_{pr}}(T_c - T_r) \quad (18)$$

Table 1: Parameters for the van de vusse reactor

Parameters	Values	Parameters	Values
k_{01} (hr^{-1})	1.287×10^{12}	c_{pc} (kJ/kg•K)	2.0000
k_{02} (hr^{-1})	1.287×10^{12}	$\tilde{n}_r$ (kg L^{-1})	0.9342
k_{03} (hr^{-1})	9.043×10	q_r (L h^{-1})	140.1900
E_1/R (*K)	9758.3	Q_c (kJ h^{-1})	-1113.5000
E_2/R (K)	9758.3	U (kJ h^{-1} m^{-2} K^{-1})	4032.0000
E_3/R (K)	8560	A_r (m^2)	0.2150
h_1 (kJ mol^{-1})	-4.2	c_{A0} (mol L^{-1})	5.1000
h_2 (kJ mol^{-1})	11	c_{B0} (mol L^{-1})	0.0000
h_3 (kJ mol^{-1})	41.85	T_{r0} (K)	387.0500
V_r (L)	10	m_c (kg)	5.0000
c_{pr} (kJ kg^{-1} K^{-1})	3.01		

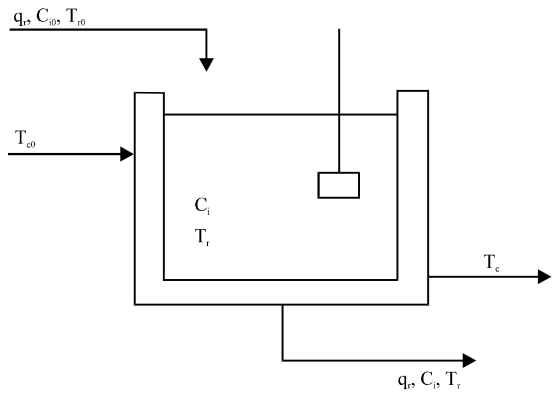

Fig. 2: The van de Vusse reactor

$$\frac{dT_c}{dt} = \frac{1}{m_c c_{pc}}\left(Q_c + A_r U (T_r - T_c)\right) \quad (19)$$

The net heat of reaction (ΔH_r) for the above reactions is expressed as:

$$\Delta H_r = \Delta h_1 k_1 c_A + \Delta h_2 k_2 c_B + \Delta h_3 k_3 c_A^2 \quad (20)$$

where, Δh_i refers to individual heat of reactions. Nonlinearity can be found in reaction rates (k_j) which are described via the Arrhenius expression:

$$k_j(T_r) = k_{0,j}\exp\left(\frac{-E_j}{RT_r}\right),\ j=1,2,3 \quad (21)$$

where, $k_{0,j}$ represents the pre-exponential factors and E_j are activation energies. Fixed parameters of the system given by Vojtesek *et al.* (2004) are shown in Table 1.

The nonlinear system identification is carried out for single-input single-output system by considering the dynamic characteristics from the changes in the space velocity, $F_v = q_r/V_r$ (h^{-1}) and the product outlet concentration, C_B(mol L^{-1}).

Multi-step ahead prediction comparison analysis: Figure 3 shows the input-output data set for the van de Vusse reactor for the model development phase (k<6000) and the out-of-sample test set (k>6000) for the

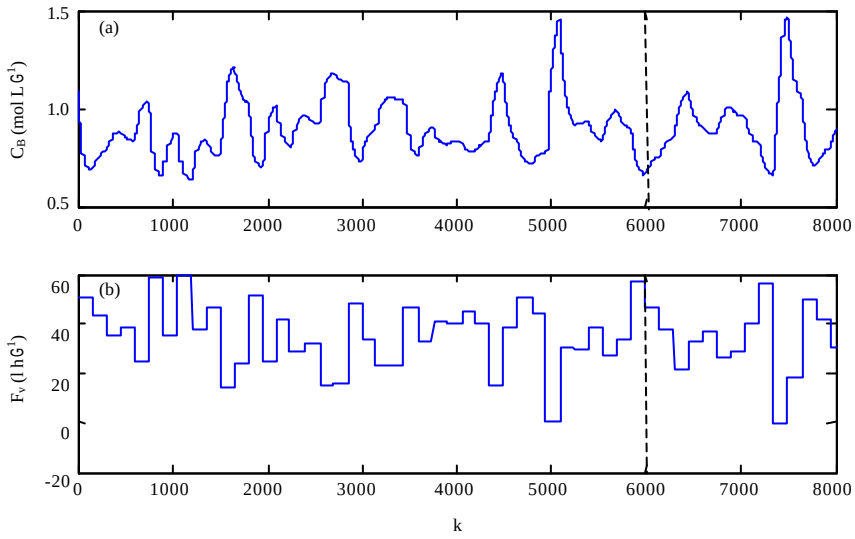

Fig. 3(a-b): Input-output data set for multi-step-ahead predictions analysis (a) F_v and (b) C_B for van de Vusse reactor

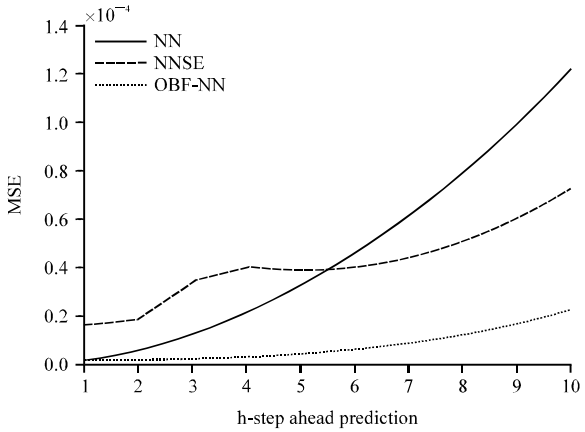

Fig. 4: Effect of prediction horizon on the model performances

multi-step-ahead prediction analysis of the proposed OBF-NN model. The estimated pole and OBF parameters using 6 Laguerre filters are:

$$\hat{p} = 0.9408$$

$$c_{OBF\text{-}NN} = [-8.5179\times10^{-4}\ -7.6897\times10^{-4}\ -2.8837\times10^{-4}\ -0.0010\ 1.7712\times10^{-4}\ -0.0014]$$

The identified residuals network for the OBF-NN model with the lowest MSE on the validation set has the following configuration:

res_{NN}: 4-5-1 neurons with tansig-linear transfer functions

The corresponding NN model with standard scaling method has 4-20-1 neurons, whereas the NN model scaled with spread encoding technique (referred to as NNSE model in this study) has 24-7-6 neurons in the input-hidden-output layers. Hyperbolic tangent sigmoid transfer functions in the hidden layer are applied to all models.

The resulting multi-step ahead prediction performances of the NN, NNSE and OBF-NN models are as shown in Fig. 4. It is observed that the prediction error for the NNSE model almost approximate a horizontal line for 3<h<8. In contrast, NN prediction capability deteriorates at a much faster rate than the other models. In this case, the OBF-NN gives the lowest multi-step ahead prediction errors for all h. Figure 5 shows the corresponding output C_B performance for 4-step ahead Fig. 5a and 7-step ahead Fig. 5b predictions. As the

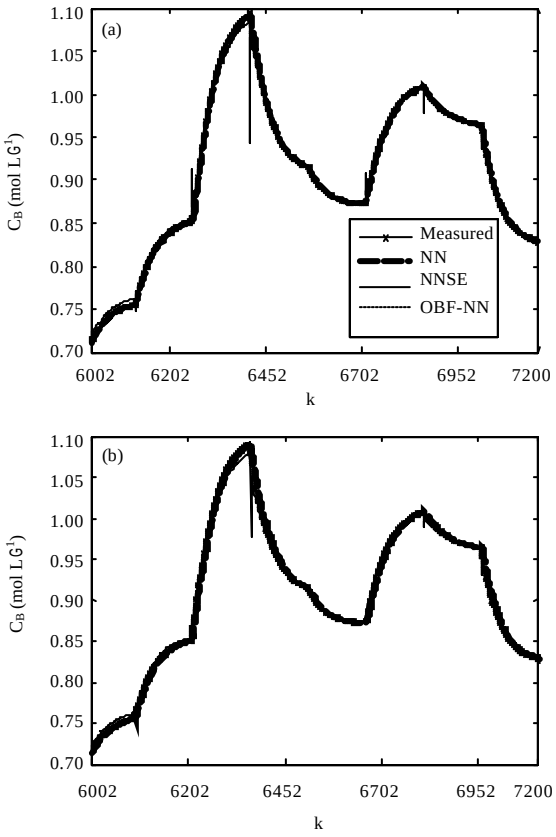

Fig. 5(a-b): Performance of NN, NNSE and OBF-NN models, (a) Four-step ahead predictions and (b) Seven-step ahead predictions

magnitude of the prediction error in Fig. 5 is small, all models seem to be able to predict the output C_B satisfactorily.

When the same models are subjected to data set in the extended region that corresponds to +14% change in the input F_v beyond the original range used to develop the models, the resulting performances differ notably as shown in Fig. 6 and 7. It is seen in Fig. 6 that the parallel OBF-NN model gives the lowest prediction error in comparison to both NN and NNSE models. It is observed that the prediction errors for both the NN and NNSE models deteriorate significantly in the extended region.

The +14% change in F_v results in the output C_B to go lower than the original limit of the training range of 0.65 mol L^{-1}. The output C_B performance for 4-step ahead and 7-step ahead predictions is shown in Fig. 7 and it is clearly seen that the parallel OBF-NN model far supersedes the other models in providing accurate predictions on the output behaviour even in the extended

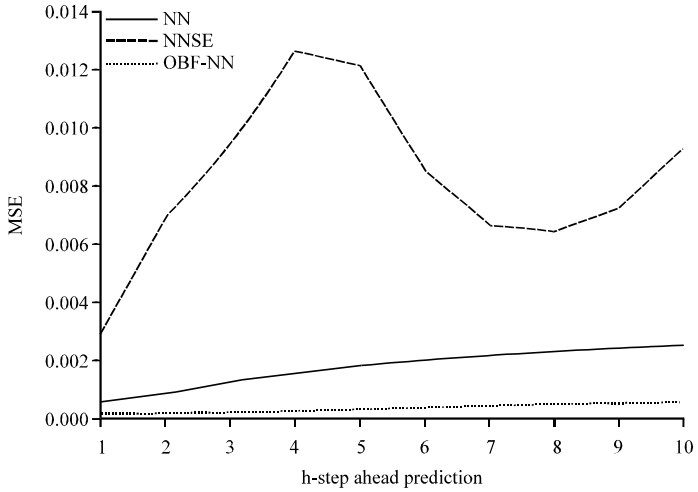

Fig. 6: Effect of prediction horizon on the model performances for data in extended region (+14% change in F_v beyond the original training range)

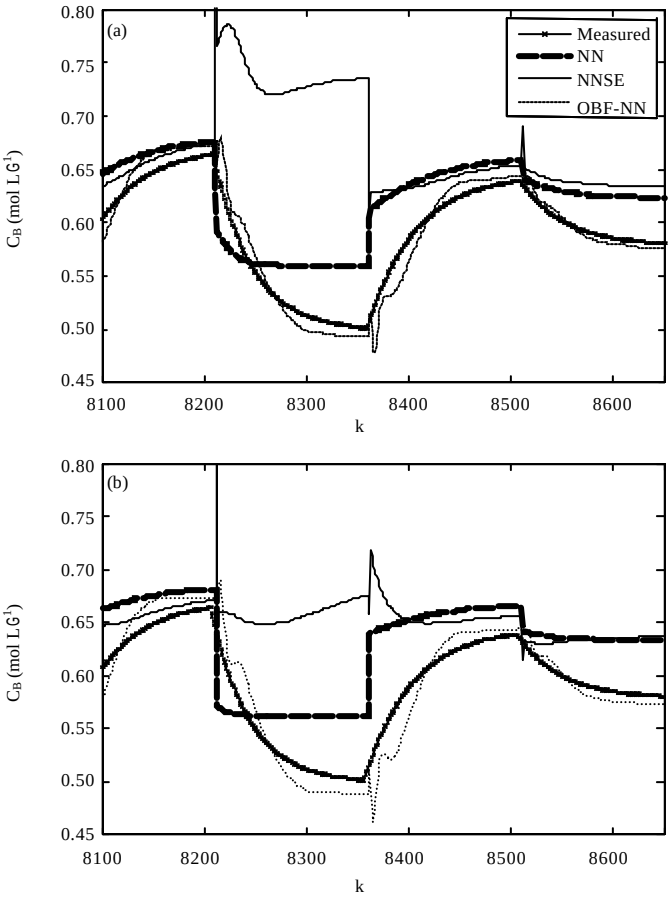

Fig. 7(a-b): Performance of NN, NNSE and OBF-NN models for data in extended region (+14% change in F_v beyond the original training range), (a) Four-step ahead predictions and (b) Seven-step ahead predictions

region. The various encode and decoding procedures inherent in the spread encoding technique renders the NNSE model unable to predict accurately in the extended region.

CONCLUSION

The current study has shown that the standard normalization procedure of scaling the data values into the range [-1,1] for NN model was adequate and provided better multi-step ahead prediction capability as the conventional NN model scaled using the spread encoding technique. Based on the results on multi-step ahead prediction analysis between OBF-NN and conventional NN scaled using standard normalization procedure as well as the spread encoding technique, it can be concluded that the OBF-NN model was able to provide satisfactory multi-step ahead prediction values similar to the conventional NN model. It can also be concluded that the OBF-NN model was superior in comparison to the other two models in providing more accurate multi-step ahead predictions in extended regions beyond the original data range used to develop the model. This shows the promising potential of embedding the parallel OBF-NN model in the framework of nonlinear MPC to improve the overall closed-loop performance over regions extended beyond the original training phase. This is certainly a promising advantage as the parallel OBF-NN model structure has the potential to allow the nonlinear MPC to be applicable to a wider operating condition. This subsequently implies lesser requirements on re-training frequency that may result in reduced maintenance cost and helps to retain the positive benefits of MPC applications in process industry.

ACKNOWLEDGMENT

Authors would like to thank Universiti Teknologi PETRONAS for funding of the project.

REFERENCES

Ardalani-Farsa, M. and S. Zolfaghari, 2010. Chaotic time series prediction with residual analysis method using hybrid Elman-NARX neural networks. Neurocomputing, 73: 2540-2553.

Cheng, H., P.N. Tan, J. Gao and J. Scripps, 2006. Multistep-ahead time series prediction. Proceedings of the 10th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining, April 9-12, 2006, Singapore, pp: 765-774.

Girard, A., C.E. Rasmussen, J. Quinonero-Candela, R. Murray-Smith, O. Winther and J. Larsen, 2002. Multiple-step ahead prediction for non linear dynamic systems-a Gaussian process treatment with propagation of the uncertainty. Adv. Neural Inform. Process. Syst., 15: 529-536.

Gomm, J.B., D. Williams, J.T. Evans, S.K. Doherty and P.J.G. Lisboa, 1996. Enhancing the non-linear modelling capabilities of MLP neural networks using spread encoding. Fuzzy Sets Syst., 79: 113-126.

Nelles, O., 2001. Nonlinear System Identification: From Classical Approaches to Neural Networks And Fuzzy Models. Springer, ISBN: 9783540673699 Berlin Pages: 785.

Piche, S., B. Sayyar-Rodsari, D. Johnson and M. Gerules, 2000. Nonlinear model predictive control using neural networks. IEEE Control Syst., 20: 53-62.

Sjoberg, J., Q. Zhang, L. Ljung, A. Benveniste and B. Delyon *et al*., 1995. Nonlinear black-box modeling in system identification: A unified overview. Automatica, 31: 1691-1724.

Tufa, L.D. and M. Ramasamy, 2009. Control relevant system identification using orthonormal basis filter models. Chemical Engineering Department, Universiti Teknologi PETRONAS, Tronoh, Malaysia.

Vojtesek, J., P. Dostal and R. Haber, 2004. Simulation and control of a continuous stirred tank reactor. Proceedings of 6th Portuguese Conference on Automatic Control, June 7-9, 2004, Faro, Portugal, pp: 315-332.

Wakasugi, H. and K. Sueyoshi, 2010. Energy-saving solutions by Advanced Process Control (APC) technology. Yokogawa Tech. Rep. English Edn., 53: 11-14.

Yasui, T., A. Moran and M. Hayase, 1996. Integration of linear systems and neural networks for identification and control of nonlinear systems. Proceedings of the 35th SICE Annual Conference on International Session Papers, July 24-26, 1996, Tottori, Japan, pp: 1389-1394.

Zabiri, H., M. Ramasamy, L.D. Tufa and A. Maulud, 2013. Integrated OBF-NN models with enhanced extrapolation capability for nonlinear systems. J. Process Control, 23: 1562-1566.

Zhao, H., J. Guiver, R. Neelakantan and L.T. Biegler, 2001. A nonlinear industrial model predictive controller using integrated PLS and neural net state-space model. Control Eng. Pract., 9: 125-133.