Gas Processing Journal
Vol. 1 No. 2, April 2013 PP. 31-40
http://uijs.ui.ac.ir/gpj

Distillation Column Identification Using Artificial Neural Network

H. Sahraie[1], G. R. Salehi[2*], A. Ghaffari[3] and M. Amidpour[4]

[1,3,4] **Mechanical Engineering Department, K. N. Toosi University of Technology, Tehran, Iran**
[2] **Mechanical Engineering Department, Islamic Azad University, Nowshahr Branch, Iran**

Abstract: In this paper, Artificial Neural Network (ANN) was used for modeling the nonlinear structure of a debutanizer column in a refinery gas process plant. The actual input-output data of the system were measured in order to be used for system identification based on root mean square error (RMSE) minimization approach. It was shown that the designed recurrent neural network is able to precisely predict and track the response of the actual system. The comparison between the results of this paper and those of the most recent published studies as NARX model indicates the significance of the proposed approach.

Keywords: Debutanizer Column, Dynamic Neural Networks, System Identification, Modeling

1. Introduction

The refinery community has recognized the importance of process automation optimization due to its benefits in terms of both profitability and tight control in product quality. Distillation of multicomponent mixtures is one of the most common separation operations in the chemical industry and refineries. Distillation is a multivariable constrained, coupled, nonlinear and nonstationary process with a complex dynamic structure. Distillation columns consume a high level of energy the accurate and tight control of which can decrease a large amount of energy consumption in refineries (Dutta & Rhinehart, 1999).

One of the problems in controlling nonlinear systems as distillation columns is thelack of an accurate model. One way of addressing this problem is to use a reliable model for the on-line prediction of the system dynamic evolution. There are essentially two approaches by which nonlinear models can be developed for a distillation column; from first principles such as trays heat and mass equations and by using the process knowledge or empirically from input/output data. The first principles using the process knowledge or empirically using input/output data. The advantages and disadvantages of each approach are well-known. In the industrial practice, it is not always possible in general to accurately obtain first models based on principle equations principles models for high-purity distillation columns. Most industrial columns are used to separate multi-component mixtures whose constituent elements are not often known completely; the fundamental thermodynamics of multi-component vapor-liquid equilibrium,

* **Corresponding Author.**

Authors' Email Address:
Hamed Sahraei (hamed.sahraie@yahoo.com), **Gholam Reza Salehi** (rezasalehi20@gmail.com), **Ali Ghaffari** (ghaffari@kntu.ac.ir), **Majid Amidpour** (amidpour@kntu.ac.ir)

ISSN (Online): 2345-4172, **ISSN (Print):** 2322-3251

the physical property data and other essential constitutive relations required for the successful development of a model based on principle equations a first principles model are not always available. And even when such knowledge is available, the resulting models usually occur in the form of a very large system of coupled nonlinear ordinary differential equations; therefore, they might not always be the most convenient ones for the controller or fault detection (FD) design. On the other hand, when fundamental process knowledge is unavailable or incomplete, or when the resulting model may not be particularly suitable for FD applications, the input/output models identified from plant data may be more useful (Chetouani, 2007).

The theory of system identification plays a significant role in many fields of science and engineering including simulation, automatic control, fault tolerant analysis, prediction, etc. (Billings* & Wei, 2005; Chen & Billings, 1992). Black box Modeling is an elegant system identification technique which uses input--output data of the process for modeling and does not need any a priori knowledge on the physical insight of the process. Many linear system identification and/or function approximation methods have been developed in the past decades (Lennart, 1999; Wahlberg, 1991). Over the years, many successful nonlinear modeling techniques have been developed such as Nonlinear Auto Regressive Moving Average with exogenous input (NARMAX) model (Billings* & Wei, 2005; Wahlberg, 1991), artificial neural networks (Lennart, 1999) and fuzzy-logic based models (Emami, Goldenberg, & Türkşen, 2000; Lennart, 1999) some combinations of which like neuro-fuzzy models (Babuška & Verbruggen, 2003) support vector machine and kernel methods of modeling (Lennart, 1999). The successful use of polynomial models (NARMAX) for nonlinear system identification has been demonstrated elsewhere (Billings* & Wei, 2005; Glass, 1999). These polynomial models can gather global information of the system dynamics quite efficiently while local information could not be approximated parsimoniously. Neural networks are another successful class of nonlinear models that have been widely used for nonlinear system modeling over the years (Chen & Billings, 1992; Nelles, 2001). Although neural network is capable of efficient identification of the process dynamics, there is no structured construction procedure for determining the number of neurons and hidden layers in the network model.

In addition to more classical identification methods such as NARMAX modeling (Korenberg & Paarmann, 1991), a new set of methods has been developed recently which apply artificial neural networks to the tasks of identification and control of dynamic systems. These works are supported by two of the most important capabilities of neural nets: their ability to learn (based on the optimization of an appropriate error function) and their good performance for the approximation of nonlinear functions (Girosi & Poggio, 1989; Narendra & Parthasarathy, 1990; Polycarpou & Ioannou, 1991).

At present, most of the works on system identification using neural nets are based on multilayer Feedforward neural networks with backpropagation learning or more efficient variations of this algorithm. These methods have been applied to real processes and they have shown an adequate behavior. It is important to remark that most of them use static discrete-time models that capture the dynamics of the real process using tapped-delay lines in the model inputs and outputs (Miller, Sutton, & Werbos; Narendra & Parthasarathy, 1990). A number of drawbacks associated with this type of models may appear in the identification of complex dynamic systems which include difficulties in selecting the appropriate number of required delays and, in some cases, poor identification performance when implemented on-line after training off-line due to training deficiencies.

Dynamic networks are generally more powerful than static ones (although somewhat more difficult to train). Because of having memory, they can be trained to learn sequential or time-varying patterns. This has applications in such disparate areas as prediction in financial markets (Roman & Jameel, 1996), channel equalization in communication systems (Feng, Tse, & Lau, 2003), phase detection in power systems (Kamwa, Grondin, Sood, Gagnon, & Mereb, 1996), sorting (Rahman, 2004), fault detection (Gan & Danai, 1999), speech recognition (Robinson, 1994), and even prediction of protein structure in genetics (Pollastri, Przybylski, Rost, & Baldi, 2002). A discussion of many more dynamic network applications can be found in Medsker and Jain (2000).

Engell and Fernholz (2003) used a semi-batch reactive distillation process. A comparison was carried out between conventional control structures and model-based predictive control using a neural net plant model. Brizuela, Uria, and Lamanna (1996) used a nonlinear model of the process for predicting future outputs that use a

Feedforward neural network (FNN). Yu and Li (2001) obtained some new results on system identification with dynamic neural networks and concluded that the gradient descent algorithm for weight adjustment was stable in an L∞ sense and robust to any bounded uncertainties. Shurong and Feng (2000) used dynamic neural network for learning input-output behaviors of a binary distillation column by combining the mechanistic property. The convergence of the algorithm was discussed using the Lyapunov method. Based on the identified model, a nonlinear adaptive controller was designed, which could preserve the stability and robustness of the closed loop system. Calderon, Draye, Pavisic, Teran, and Libert (1996) worked with the Dynamical Recurrent Neural Network as a tool for system identification and trained the network using a time-dependent back propagation learning algorithm and showed that, for modeling a nonlinear dynamical system, their neural device had good performance for interpolation and extrapolation and was very robust in the presence of noise.

Tavakolmoghadam and Safavi (2012) introduced an experimental based ANN model to describe the performance of vacuum membrane distillation process for desalination in different operating conditions. Tonnang and Olatunbosun (2010) developed NN that could be used for the control of an industrial process. Field data from a working distillation column or fractionator of crude oil refinery in Nigeria was used for the development and testing the effectiveness of the controller. Sharma and Singh (2012) developed three different control strategies, viz., conventional PID control, model predictive control and neural network predictive control that are implemented to a reactive distillation column. Konakom, Kittisupakorn, and Mujtaba (2012) developed neural network-based model predictive control incorporated with a neural network estimator for handling the predefined optimal policy tracking of a batch reactive distillation. de Canete, Gonzalez-Perez, and del Saz-Orozco (2008) utilized LABVIEW platform to provide a powerful toolset for process identification and control of nonlinear systems based on artificial neural networks. This tool has been applied to the monitoring and control of a lab-scale distillation column.

In this paper, the real distillation column was identified and modeled with two approaches: first, with ARX model; then with the recurrent dynamic network with feedback connections enclosing several layers of the network. Finally, the results were compared with each other.

2. Distillation Column Dynamic Modeling and Simulation

In this section, mathematical modeling of distillation column is introduced. The column for which the model was presented separated a single multicomponent liquid feed into two liquid products in a tray-type distillation column. The column was equipped with a reboiler and a total condenser. In mathematical modeling, some simplifying assumptions are usually considered. The model assumed that vapor holdups were negligible and that the effluent streams were in thermodynamic equilibrium. The column pressure was assumed to remain constant throughout the dynamic tests. The dynamic of the reboiler and the condenser were neglected. Finally, the dynamic changes in internal energy on the trays were assumed to be so rapid that the energy equation reduced to an algebraic equation. With the foregoing assumptions, the dynamic model can be expressed by the following set of differential and algebraic equations.

Overall mass balance for each tray:

$$\frac{dM_n}{dt} = L_{n+1} + V_{n-1} - L_n - V_n \quad (1)$$

Component balance for each tray:

$$\frac{dM_n X_{n,j}}{dt} = L_{n+1}X_{n+1} + V_{n-1}y_{n-1,j} - L_n X_{n,j} - V_n y_{n,j} \quad (2)$$

Energy balance for tray n:

$$L_{n+1}h_{n+1} + V_{n-1}yh_{n-1} - L_n h_n - V_n h_n = 0 \quad (3)$$

Tray hydraulics:
If the Francis Weir formula was used, the relationship would be:

$$L_n = 3.33\, l h_n^{\frac{3}{2}} \quad (4)$$

where l is the length of weir in m, h_n is the height of liquid over weir in m and Ln is the liquid leaving stage n, m^3/sec.
Phase equilibrium:

$$y_{n,j} = K_{n,j}x_{n,j} \quad (5)$$

Murphree vapor-phase efficiency:

$$y^A_{n,j} = y^A_{n-1,j} + \eta_{ij}(y_{n,j} - y^A_{n-1,j}) \quad (6)$$

where superscript A denotes actual concentration (Deshpande & Plank, 1985).

3. Dynamic Neural Networks

Multilayer Perceptrons (MLP) allow only for feed forward connections between each neuron and the neurons in the following layer. Recurrent neural networks in contrast allow for arbitrary connections between neurons, both forward and recurrent (feedback). A nonlinear mapping obtained by a recurrent

neural network depends not only on the current input, but also on the previous inputs by the feedback connections to the input.

The recurrent neural network used in this work was a special class of recurrent networks where recurrent connections were only allowed within the nodes of the output layer to input layer. A simplified schematic of recurrent dynamic network is shown in Fig. 1. Training data was scaled in the range of [0, 1] in order to avoid saturating the nonlinearities of the nodes' activation functions.

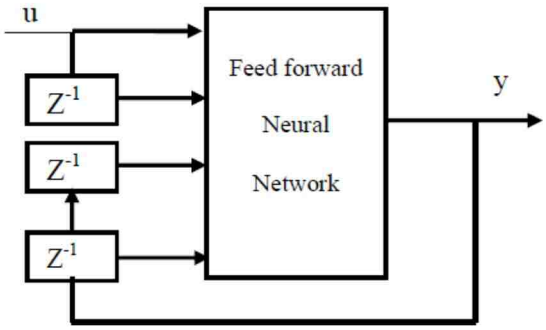

Figure 1. Recurrent dynamic network schematic

3.1. Dynamic Network Structures

A dynamic network has memory. Its response at any given time depends not only on the current input, but on the history of the input sequence. If the network does not have any feedback connections, then only a finite amount of history will affect the response.

Each layer in the RNN is made up of the following parts:

- A set of weight matrices that come into that layer (which can connect from other layers or from external inputs), associated weight function rule used to combine the weight matrix with its input (normally standard matrix multiplication) and associated tapped delay line
- Bias vector
- Net input function rule that is used to combine the outputs of various weight functions with the bias to produce the net input (normally a summing junction)
- Transfer or activation function

P1 is the first input to first layer and Z-1 is time delay. The network has inputs that are connected to special weights, called input weights, and are denoted by IWi,j where j denotes the number of the input vector that enters the weight, and i denotes the number of the layer to which the weight is connected. The weights connecting one layer to another are called layer weights and are denoted by LWi,j where j denotes the number of the layer coming into the weight, and i denotes the number of the layer at the output of the weight. f, g are transfer or activation functions (e.g. tangent hyperbolic, threshold function and sigmoid function).

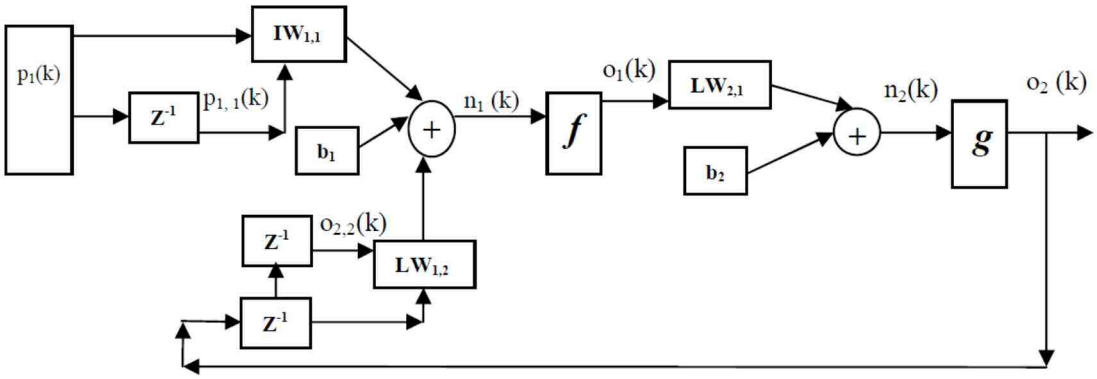

Figure 2. Dynamic multilayer network

3.2. Dynamic Network Training

Although dynamic networks can be trained using the same gradient-based algorithms that are used for static networks, the performance of the algorithms on dynamic networks can be quite different, and the gradient must be computed in a more complex way.

Consider the multilayer network in Fig. 2. The basic simulation equation of such a network is

$$o_1(k) = f\left(IW_{1,1}\left(p_1(k);\ p_{1,1}(k)\right) + LW_{1,2}o_{2,2}(k) + b_1\right) = f\left(n_1(k)\right) \quad (7)$$

$$o_2(k) = g\left(LW_{2,1}o_1(k) + b_2\right) = g\left(n_2(k)\right) \quad (8)$$

The generalized error expression includes all squared errors at outputs k=1,2, …,K.

$$E = \frac{1}{2}\sum_{k=1}^{k}(d_k - o_2(k))^2 \quad (9)$$

Assume that the gradient decent search is performed to reduce the error E through the adjustment of weights. Requiring the weight adjustment, individual weight adjustment can be computed as follows:

$$\Delta LW_{2,1} = -\eta\frac{\partial E}{\partial LW_{2,1}} \quad (10)$$

where η is learning rate. Define

$$S_2(k) \equiv \frac{\partial E}{\partial n_2(k)} \quad (11)$$

Using the chain rule, that the following can be shown:

$$\frac{\partial E}{\partial LW_{2,1}} = \frac{\partial E}{\partial n_2(k)}\frac{\partial n_2(k)}{\partial LW_{2,1}} \quad (12)$$

It is known that:

$$\frac{\partial n_2(k)}{\partial LW_{2,1}} = o_1(k) \quad (13)$$

And, according to Eqs. (8) And (9), the first term of right hand side in Eq. (12) is obtained:

$$S_2(k) \equiv \frac{\partial E}{\partial n_2(k)} = \frac{\partial}{\partial n_2(k)}\left\{\frac{1}{2}\sum_{k=1}^{k}(dk - g(n_2(k))^2\right\} = -\sum_{k=1}^{k} g'(n_2(k))(dk - o_2(k)) \quad (14)$$

Totally:

$$\Delta LW_{2,1} = \eta \times S_2(k) \times o_1(k)^T \quad (15)$$

And, for hidden layers:

$$\Delta LW_{1,2} = -\eta\frac{\partial E}{\partial LW_{1,2}} \quad (16)$$

$$S_1(k) \equiv \frac{\partial E}{\partial n_1(k)} \quad (17)$$

$$\frac{\partial E}{\partial LW_{1,2}} = \frac{\partial E}{\partial n_1(k)}\frac{\partial n_1(k)}{\partial LW_{1,2}} \quad (18)$$

From Eq. (7), the following relation can be obtained:

$$\frac{\partial n_1(k)}{\partial LW_{1,2}} = o_{2,2}(k) \quad (19)$$

$$S_1(k) \equiv \frac{\partial E}{\partial n_1(k)} = \frac{\partial E}{\partial f}\frac{\partial f}{\partial n_1(k)} = \left(\frac{\partial}{\partial f}\left\{\frac{1}{2}\sum_{k=1}^{k}(dk - g(n_2(k))^2\right\}\right) \times f'(n_1(k)) \quad (20)$$

$$\frac{\partial}{\partial f}\left\{\frac{1}{2}\sum_{k=1}^{k}(dk - g(n_2(k))^2\right\} \quad (21)$$

$$= -\sum_{j=1}^{j}\sum_{k=1}^{k}((dk - o_2(k)) \times g'(n_2(k)) \times LW_{2,1}(i,j))$$

And, totally, for a hidden layer, the following relation can be presented:

$$\Delta LW_{1,2} = \eta \times S_1(k) \times o_{2,2}(k)^T \quad (22)$$

For other weight matrices of other hidden layers and biases, this approach can be employed.

4. Description of the Plant

The column was located in a refinery in Iran, and it was a part of naphtha splitter plant. In the debutanizer column, C3 (propane) and C4 (butane) were removed from the heavier compositions such as C5 (pentane).

The tasks of debutanizer column were as below;

- Preparing sufficient fractionation
- Maximizing the C5 (stabilized gasoline) content in the distillate of debutanizer while respecting the limit enforced by law
- Minimizing the C4 (butane) content in the debutanizer bottoms

A detailed scheme of the debutanizer column is shown in Fig. 3. A number of sensors were installed on the plant to monitor product quality. The subsets of sensors relevant to the described application together with the corresponding description are listed in Table 1.

Table 1. Sensors relevant to the described application and corresponding characteristics

Tag	Description	Units
TI 6001	Feed temperature	° C
FI 6000	Feed flow	Kbbl / day
TI 6002	Bottom temperature	° C
TI 6006	Top temperature	° C
PI 6006	Top pressure	bar
FI 6002	Reflux flow	m^3 / hr
FI 6001	Steam flow	m^3 / hr
G.C	Gas Chromatograph	mole fraction

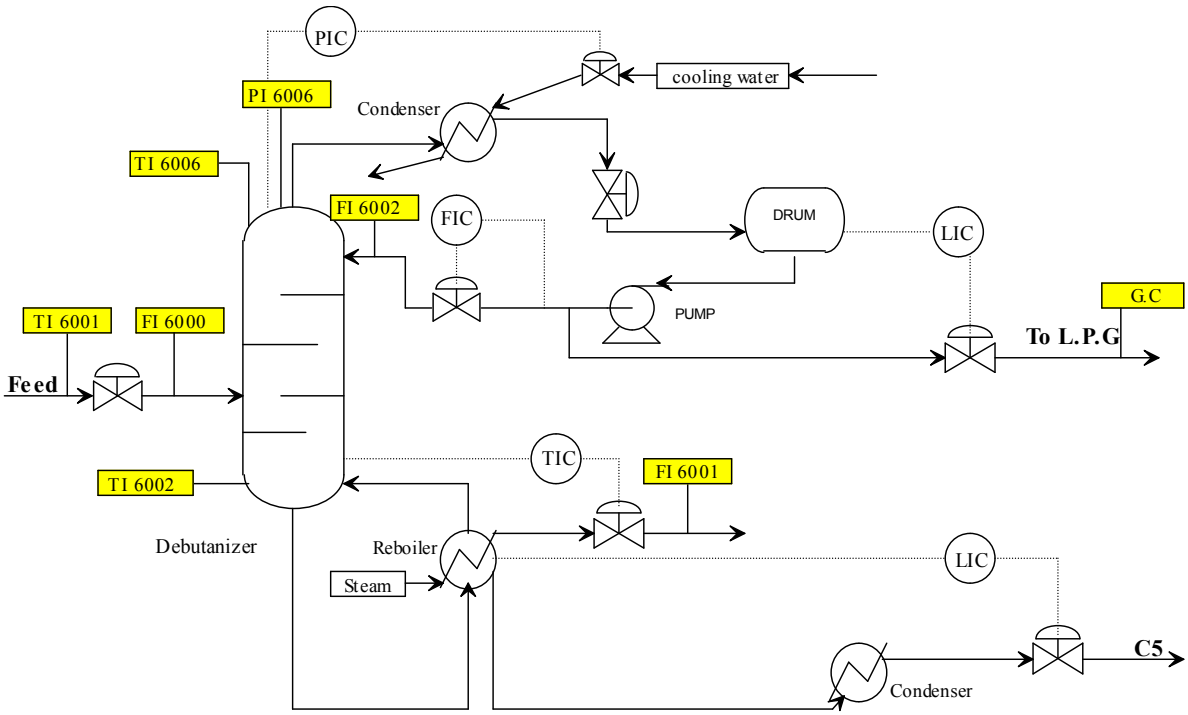

Figure 3. The schematic diagram of the debutanizer column

5. Simulation Results

As shown in Fig.3, there were five control loops in the debutanizer column; so, there were five controlled and five manipulated variables. Another variable that changed during the column operation was the feed flow. The model was structured based on inputs which manipulated variables in control loops (reflux flow, steam flow entering to reboiler, cooling water flow entering to condenser, top and bottom products flow) plus feed flow variable and outputs which were control variables (top and bottom temperature, pressure of column, level of drum liquid, level of reboiler liquid). The model output can be expressed as:

$$y(k) = f(y(k-1), y(k-2), u(k), u(k-1)) \quad (23)$$

Where u is the vector of input variables and y is output variables. The number of time delay of input and output was achieved by trial and error method. About 340 pairs of data were used for this model around 70% and 30% of which were employed for training and validation, respectively. Before training and validation, all data were normalized and put in the range of [0,1].

At first, nonlinear ARX model (Lennart, 1999; Ramesh, Aziz, & Shukor, 2008) was constructed based on the above named variables. Five MISO models were made and, for each one, about twenty models were examined; the models were compared with each other according to root mean square error (RMSE) criterion. The root mean square error is defined as

$$RMSE = \sqrt{\frac{1}{N}\sum_{p=1}(t_p - o_p)^2} \quad (24)$$

In Eq. (24), N represents the number of input/output training pairs, t_p is the target output for the p-th training and op is the simulated output. Then, recurrent ANN model was structured with variables used for NARX model. Although the ANN model can be constructed based on six inputs and five outputs, since the network with five outputs was so complicated and selecting the optimal net based on RMSE criterion was very difficult, five networks with one output were made.

The performance of network depends on the architecture of network as the number of layers and nodes. To select an optimal network, the root mean square error criterion was used. The number of hidden nodes was obtained by a trial and error method. Table 2 shows the minimum value of RMSE of recurrent ANN obtained by the trial and error process. Fig.4 shows the real and simulated outputs of ANN and NARX models by training data. However, to obtain a general model, a net needed to be selected on the basis of validation data and RMSE of validation data. Fig.5 shows the real and simulated outputs of ANN and NARX models by validation data, and the absolute error between real and simulated outputs of ANN model in validation data is shown in Fig.6.

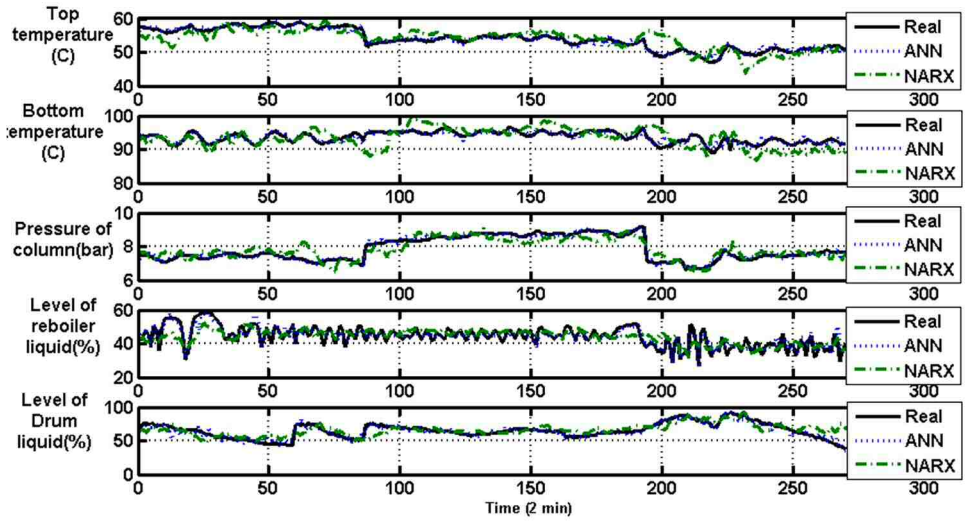

Figure 4. Real and simulated outputs of training data

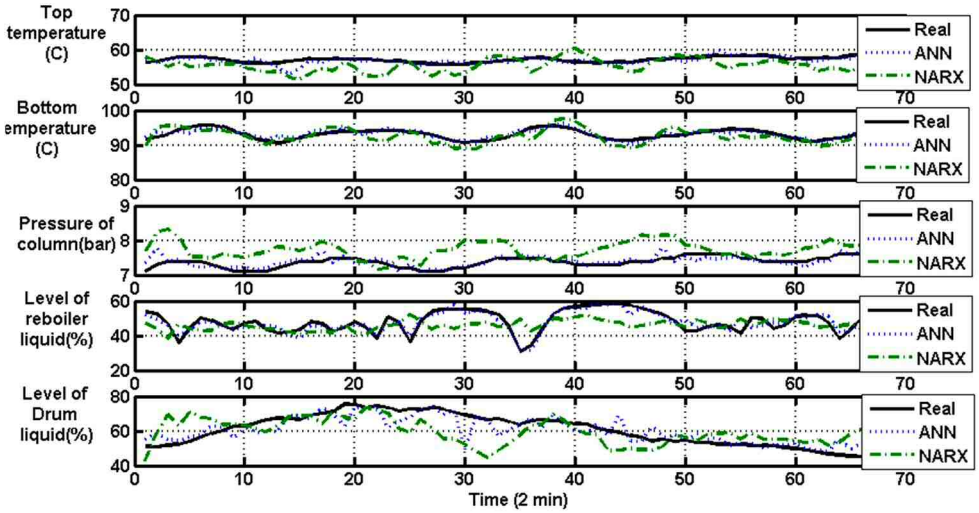

Figure 5. Real and simulated outputs of validation data

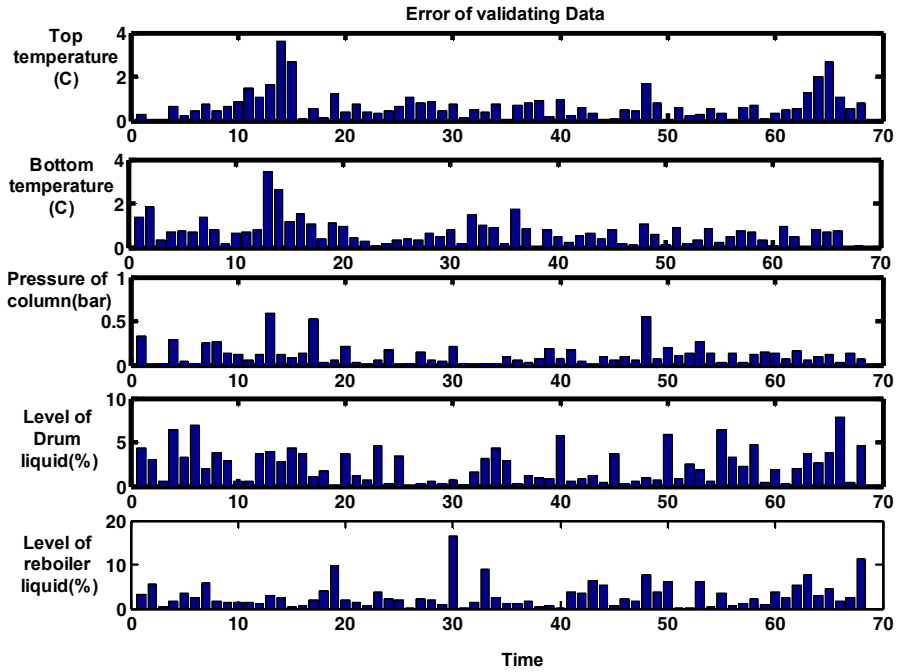

Figure 6. Absolute error between real and simulated outputs of ANN model with validation data

Table 2. Configuration and minimum RMSE of networks

Output variable	Network configuration	Minimum RMSE
Top temperature of column (°C)	6_10_1	0.9562
Bottom temperature of column (°C)	6_4_1	0.7548
Pressure of column (bar)	6_11_1	0.1397
Level of reboiler liquid (percent of total height)	6_9_1	2.8527
Level of drum liquid (percent of total height)	6_13_1	4.5656

Table 3. Comparison of ANN model and NARX model with training data

	Neural Network model				NARX model			
	RMSE	MRE	R^2	STD	RMSE	MRE	R^2	STD
Top Temp	0.6889	0.549	0.9452	0.6889	2.066	1.5832	0.782	1.6957
Bottom Temp	0.7361	0.651	0.8575	0.7895	2.509	2.009	0.658	1.7279
Pressure	0.1285	0.0961	0.968	0.0286	0.3616	0.1307	0.727	0.2655
Reboiler level	2.8169	2.1467	0.8475	2.8246	4.5627	3.5817	0.602	4.9288
Drum level	4.0481	3.5874	0.8719	4.2671	8.1198	6.0672	0.589	8.2157

Table 4. Comparison of ANN model and NARX model with validation data

	Neural Network model				NARX model			
	RMSE	**MRE**	**R^2**	**STD**	**RMSE**	**MRE**	**R^2**	**STD**
Top Temp	0.9562	0.836	0.754	1.02	2.5845	2.2125	0.528	3.005
Bottom Temp	0.7548	0.921	0.7557	1.3612	2.5978	2.3423	0.508	2.34
Pressure	0.1397	0.267	0.873	0.154	0.4276	0.3566	0.636	0.6412
Reboiler level	2.8527	2.674	0.7789	3.021	6.7158	5.466	0.534	5.0324
Drum level	4.5656	4.0011	0.7823	5.0851	9.2112	7.3871	0.4648	9.7372

In addition to RMSE criterion, there are some other criteria to compare the performance of two models with each other and with the real data. The performance of ANN and NARX configurations are compared using the mean relative error (MRE) and the standard deviations in the relative (STD) error and the coefficient of determination, R2, of the linear regression line between the predicted values from the models and the desired output. These three criteria can be defined as:

$$MRE = \frac{1}{N}\sum_{p=1}^{N} \quad S(E) \tag{25}$$

$$STD = \sqrt{\frac{\sum_{p=1}^{N}(E \ -\bar{E})^2}{N-1}} \tag{26}$$

$$R^2 = 1 - \frac{\sum_{p=1}^{N}(t_p - o_p)^2}{\sum_{p=1}^{N}(o_p)^2} \tag{27}$$

Where Er = (tp-op)/op. N represents the number of input/output training pairs, tp is the target output for the p-th training and op is the simulated output.

These criteria were calculated for both ANN and NARX models with training and validation data and the results are reported in Tables 3 & 4. Since the information of validation data was based on selecting optimum model as shown in Table 4, the ANN model can be used for more accurate identification of a nonlinear system as a better approach.

6. Conclusion

In this paper, recurrent neural network was implemented for modeling and identifying a debutanizer distillation column. This model was based on real data of a distillation in one of the Iranian refineries. The configuration of neural nets was achieved using a trial and error approach. The RMSE was a criterion for selecting the optimal networks; according to Table 2, RMSEs were acceptable for top temperature, bottom temperature, pressure, level of reboiler and drum liquid and ANN model for the aim of control. Then, NARX model was compared with recurrent ANN model with some criteria such as RMSE, MRE, STD and R2. The results show that ANN model was more capable and accurate than NARX model. Because ANN can accept feedback from all layers and also it has high flexibility in structure (to be added more nodes, layers), the accurate and better model can be obtained.

Nomenclature

f,g	transfer or activation function
Δw	weight increment for the input layer of neurons
η	learning parameter (positive constant)
ηij	Murphree stage efficiency
B	bottom product rate (kmols/h)
d_k	desired output vector for kth pattern
D	distillate product rate (kmols/h)
E	least squared error for pth pattern
Fi	total feed flow rate into ith tray (kmols/h)
h_F	total molar enthalpy of feed (kJ/kmol)
h_{fij}	component feed enthalpy (kJ/kmol)
h_i	total molar enthalpy of liquid mixture (kJ/kmol)
H_i	total molar enthalpy of vapor (kJ/kmol)
h_{lij}	component liquid enthalpy (kJ/kmol)
H_{vij}	component vapor enthalpy (kJ/kmol)
K_{ij}	equilibrium constant
L_i	total liquid flow rate leaving the tray (kmols/h)
L_{ij}	component liquid flow rate leaving the ith tray (kmols/h)
M_B	liquid molar holdup in reboiler (kmols)
M_D	liquid molar in reflux drum
M_i	liquid molar holdup on ith tray (kmols)
NC	number of components
v	scalar product of weight vector and input vector
v_I	scalar product of ith weight vector and input vector
NT	total number of trays in distillation column
o	output vector of neuron
o_k	kth output of neurons processing nodes
Q_B	reboiler heat duty (KJ/h)
Q_C	condenser heat duty (KJ/h)
R	reflux rate (kmols/h)
U	side stream in liquid phase
Vi	total vapor flow rate from the tray (kmols/h)
v_{ij}	component vapor flow rate from the tray (kmols/h)
W	side stream in vapor phase
x	liquid composition of more volatile component (mole fraction)
x_{Fij}	component liquid composition of jth component in feed (mole fraction)
x_{ij}	liquid composition of jth component on ith tray (mole fraction)
y	vapor composition of more volatile component (mole fraction)
y^*	equilibrium vapor composition of more volatile component (mole fraction)
y_{ij}	vapor composition of jth component on ith tray (mole fraction)
y_{ij}^*	equilibrium vapor composition of jth component on ith tray (mole fraction)
yn	input vector to neuron layer
ANN	Artificial Neural Network

References

Babuška, R., & Verbruggen, H. (2003). Neuro-fuzzy methods for nonlinear system identification. *Annual reviews in control, 27*(1), 73-85.

Billings*, S. A., & Wei, H.-L. (2005). The wavelet-NARMAX representation: a hybrid model structure combining polynomial models with multiresolution wavelet decompositions. *International Journal of Systems Science, 36*(3), 137-152.

Brizuela, E., Uria, M., & Lamanna, R. (1996). *Predictive control of a multi-component distillation column based on neural networks.* Paper presented at the Neural Networks for Identification, Control, Robotics, and Signal/Image Processing, 1996. Proceedings., International Workshop on.

Calderon, G., Draye, J.-P., Pavisic, D., Teran, R., & Libert, G. (1996). *Nonlinear dynamic system identification with dynamic recurrent neural networks*. Paper presented at the Proc. of the International Workshop on Neural Networks for Identification, Control, Robotics and Signal/Image Processing.

Chen, S., & Billings, S. (1992). Neural networks for nonlinear dynamic system modelling and identification. *International Journal of Control, 56*(2), 319-346.

Chetouani, Y. (2007). Using Artificial Neural networks for the modelling of a distillation column. *IJCSA, 4*(3), 119-133.

de Canete, J. F., Gonzalez-Perez, S., & del Saz-Orozco, P. (2008). Artificial Neural Networks for Identification and Control of a Lab-Scale Distillation Column using LABVIEW. *International Journal of Intelligent Technology, 3*(2).

Deshpande, P. B., & Plank, C. A. (1985). *Distillation dynamics and control*: Instrument Society of America.

Dutta, P., & Rhinehart, R. R. (1999). Application of neural network control to distillation and an experimental comparison with other advanced controllers. *ISA transactions, 38*(3), 251-278.

Emami, M. R., Goldenberg, A. A., & Türkşen, İ. B. (2000). Systematic design and analysis of fuzzy-logic control and application to robotics: Part I. Modeling. *Robotics and Autonomous Systems, 33*(2), 65-88.

Engell, S., & Fernholz, G. (2003). Control of a reactive separation process. *Chemical Engineering and Processing: Process Intensification, 42*(3), 201-210.

Feng, J., Tse, C. K., & Lau, F. C.-m. (2003). A neutral-network-based channel-equalization strategy for chaos-based communication systems. *IEEE transactions on circuits and systems. I, Fundamental theory and applications, 50*(7), 954-957.

Gan, C., & Danai, K. (1999). *Fault diagnosis of the IFAC Benchmark Problem with a model-based recurrent neural network.* Paper presented at the Control Applications, 1999. Proceedings of the 1999 IEEE International Conference on.

Girosi, F., & Poggio, T. (1989). Representation properties of networks: Kolmogorov's theorem is irrelevant. *Neural Computation, 1*(4), 465-469.

Glass, J. W. (1999). NARMAX modelling and robust control of internal combustion engines. *International Journal of Control, 72*(4), 289-304.

Kamwa, I., Grondin, R., Sood, V., Gagnon, C., & Mereb, J. (1996). Recurrent neural networks for phasor detection and adaptive identification in power system control and protection. *Instrumentation and Measurement, IEEE Transactions on, 45*(2), 657-664.

Konakom, K., Kittisupakorn, P., & Mujtaba, I. M. (2012). Neural network-based controller design of a batch reactive distillation column under uncertainty. *Asia-Pacific Journal of Chemical Engineering, 7*(3), 361-377.

Korenberg, M. J., & Paarmann, L. D. (1991). Orthogonal approaches to time-series analysis and system identification. *Signal Processing Magazine, IEEE, 8*(3), 29-43.

Lennart, L. (1999). System identification: theory for the user: Prentice Hall PTR, USA.

Medsker, L., & Jain, L. (2000). Recurrent neural networks: design and applications.|| CRC Press. *Boca Raton, FL*.

Miller, W., Sutton, R., & Werbos, P. Neural networks for control, 1990: MIT Press.

Narendra, K. S., & Parthasarathy, K. (1990). Identification and control of dynamical systems using neural networks. *Neural Networks, IEEE Transactions on, 1*(1), 4-27.

Nelles, O. (2001). *Nonlinear system identification: from classical approaches to neural networks and fuzzy models*: Springer.

Pollastri, G., Przybylski, D., Rost, B., & Baldi, P. (2002). Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles. *Proteins: Structure, Function, and Bioinformatics, 47*(2), 228-235.

Polycarpou, M. M., & Ioannou, P. A. (1991). *Identification and control of nonlinear systems using neural network models: Design and stability analysis*: Citeseer.

Rahman, S. (2004). A neural network with O (N) neurons for ranking N numbers in O (1/N) time. *Circuits and Systems I: Regular Papers, IEEE Transactions on, 51*(10), 2044-2051.

Ramesh, K., Aziz, N., & Shukor, S. (2008). Development of NARX Model for Distillation Column and Studies on Effect of Regressors. *Journal of Applied Sciences, 8*(7).

Robinson, A. J. (1994). An application of recurrent nets to phone probability estimation. *Neural Networks, IEEE Transactions on, 5*(2), 298-305.

Roman, J., & Jameel, A. (1996). *Backpropagation and recurrent neural networks in financial analysis of multiple stock market returns.* Paper presented at the System Sciences, 1996., Proceedings of the Twenty-Ninth Hawaii International Conference on.

Sharma, N., & Singh, K. (2012). Model predictive control and neural network predictive control of TAME reactive distillation column. *Chemical Engineering and Processing: Process Intensification, 59*, 9-21.

Shurong, L., & Feng, L. (2000). *Dynamic neural network based nonlinear adaptive control for a distillation column.* Paper presented at the Intelligent Control and Automation, 2000. Proceedings of the 3rd World Congress on.

Tavakolmoghadam, M., & Safavi, M. (2012). An optimized neural network model of desalination by vacuum membrane distillation using genetic algorithm. *Procedia Engineering, 42*, 106-112.

Tonnang, H. E., & Olatunbosun, A. (2010). NEURAL NETWORK CONTROLLER FOR A CRUDE OIL DISTILLATION COLUMN. *Journal of Engineering & Applied Sciences, 5*(6).

Wahlberg, B. (1991). System identification using Laguerre models. *Automatic Control, IEEE Transactions on, 36*(5), 551-562.

Yu, W., & Li, X. (2001). Some new results on system identification with dynamic neural networks. *Neural Networks, IEEE Transactions on, 12*(2), 412-417.