ORIGINAL ARTICLE

Predicting Liquid-Vapor (LV) composition at distillation column

Totok R. Biyanto, Totok Suhartanto and Bambang L. Widjiantoro

Abstract

Biyanto,TR., Suhartanto, T. and Widjiantoro, BL.
Predicting Liquid-Vapor (LV) composition at distillation column
Songklanakarin J. Sci. Technol., 2007, 29(2) : 575-581

This paper will present the development of nonlinear model of distillation column using neural networks approach. The model is accomplished in Nonlinear Auto Regressive with exogenous input (NARX) structure. This distillation column has two input and two output variables. The input variables are heat duty on the reboiler (Qr), and reflux flowrate (L), while the output variables are mole fraction of distillate (Xd) and mole fraction bottm product (Xb). The training as well as validation data were generated using Amplitude Pseudo Random Binary Signal (APRBS) as excitation signal. The structure of neaural networks is feedforward networks, which consists of three layers: input, hidden and output layer. Levenberg-Marquardt algorithm is used as learning algorithm to adjust the weight matrices of the networks. The results show that NN soft sensor base on flow rate correlation is easy to build, fast response, no need special instrumentations, better of reliability compare to analyzer reliability, cheaper, low operational cost, low maintenance cost, and has good Root Mean Square Error (RMSE).

Key word: NN sot senspr, predicting LV-composition, distillation column

Department of Engineering Physic, FTI-ITS Surabaya Kampus, ITS Keputih, Sukolilo, Surabaya Indonesia 60111
Corresponding e-mail: trb@ep.its.ac.id
Received , 20 August 2005 Accepted, 4 August 2006

Distillation column is used to separate the binary mixture methanol-water by differences in volatilty. Thus, the system results in nonlinear and complex characteristics. Liquid-Vapor (LV) control structure is the best control pairing in binary distillation column system. In order to maintain mole fraction of distillate and bottom product, this structure will regulate the reflux flow rate & steam flow rate respectively.

The composition (in mole fraction) is generally measured by interferential measuring method. This method was chosen because of the analyzer performance which was a slow response, lack of reliability and high cost (Luyben,*et al.*1998; Zamprogna, *et al.*2000; Ruiz, *et al.* 2004)

Flow rate and mole fraction have a nonlinear relationship, which is influenced by column pressure, feed composition, feed flow rate, condenser level etc. According to these factors, using a soft sensor instrument must fulfill all the requirements, such as : is not interfered by others factors (column pressure, feed composition, feed flow rate, condenser level and etc), and is nonlinear, simple and easy to build (Biyanto,*et al.* 2004)

The objective of this work is to appliy Neural Network (NN), using Multi Layer Perception (MLP) structure (6 history lenghts, 13 hidden nodes & trained for 200 iterations), to estimate distiluate & bottom product composition (in mole fraction). Experiments were performed on a rigorous model developed in Hysys software, and NN soft sensor was implemented in Matlab.

Distillation Column

The column contains a total of NT theoretical trays. The liquid hold up on each tray including the downcomer is Mn. The liquid on each tray is assumed to be perfectly mixed with composition Xn. (Luyben,*et al.* 1990). The mathematical formula expressing what is going on in the distillation column using rigorous modeling will described as follows (Biyanto,*et al.* 2004) :

N^{th} tray

Mass balance:

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

Component mass balance:

$$\frac{d(M_n x_n)}{dt} = L_{n+1}X_{n+1} - L_n X_n + V_{n-1}Y_{n-1} - V_n Y_n \text{.....(2)}$$

Energy balance:

$$\frac{d(M_n h_n)}{dt} = L_{n+1}{}^{h}{}_{n+1} - L_n h_n + V_{n-1}H_{n-1} - V_n H_n \text{.....(3)}$$

Condenser and reflux drum

The overhead vapor is totally condensed in a condenser and flows into the reflux drum, whose holdup of liquid is MD (moles). The content of the drum is at its bubble point. Reflux is pumped back to the top tray (NT) of column at a rate L. Overhead distillate product is removed at a rate D

Mass balance:

$$\frac{dM_D}{dt} = V_{NT} - L_{NT+1} - D \quad \text{.....(4)}$$

Component mass balance:

$$\frac{d(M_D x_D)}{dt} = V_{NT}Y_{NT} - (L_{NT+1} + D)X_D \quad \text{.....(5)}$$

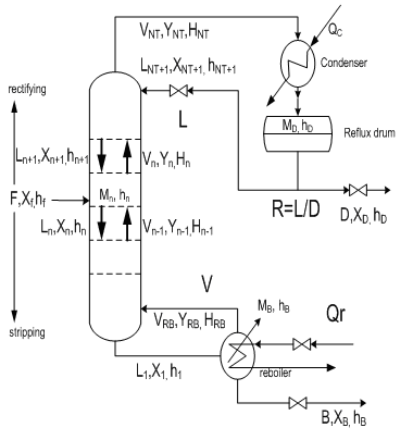


Figure 1. Binary Distillation Column Condenser and reflux drum

Energy balance:

$$\frac{d(M_D h_D)}{dt} = V_{NT}H_{NT} - L_{NT+1}H_{NT+1} - Dh_D + Q_C \quad ....(6)$$

Reboiler and base column

At the base of the column, liquid bottom product is removed at a rate B and with composition X_B. Vapor boil up is generated in the thermosiphon reboiler at rate V. Liquid circulates from the bottom of the column through the tubes in the vertical tube-in shell reboiler because of the smaller density of the vapor liquid mixture in the reboiler tubes. We will assume that the liquids in the reboiler and in the base of the column are perfectly mixed and have the same composition X_B and total holdup M_B (moles). The composition of the vapor leaving the base of the column and entering tray 1st is y_B. It is in equilibrium with the liquid with composition X_B.
Mass balance:

$$\frac{dM_n}{dt} = L_1 - V_{RB} - B \quad .....(7)$$

Component mass balance:

$$\frac{d(M_B X_B)}{dt} = L_1 X_1 - V_{RB}Y_{RB} - BX_b \quad .....(8)$$

Energy balance:

$$\frac{d(M_B h_B)}{dt} = L_1 h_1 - V_{RB}H_{RB} - Bh_b + Q_r \quad .....(9)$$

Feed tray (n = NF)

A single feed stream is fed as saturated liquid onto the feed tray N_F. Feed flow rate is F (mole/hour), enthalpies hF and composition X_F (mole fraction more volatile component), and by change n with N_F, the equations 1,2, and 3 will become equations 10, 11 and 12, respectively.
Mass balance:

$$\frac{dM_{NF}}{dt} = L_{NF+1} - L_{NF} + V_{NF-1} - V_{NF} + F \quad .....(10)$$

Component mass balance:

$$\frac{d(M_{NF}X_{NF})}{dt} = L_{NF+1}X_{NF+1} - L_{NF}X_{NF} + V_{NF-1}Y_{NF-1} - V_{NF}V_{NF} + FX_F \quad .....(11)$$

Energy balance:

$$\frac{d(M_{NF}h_{NF})}{dt} = L_{NF}X_{NF+1} - L_{NF}h_{NF} + V_{NF-1}H_{NF-1} - V_{NF}H_{NF} + Fh_F \quad ...(12)$$

Neural Networks

Neural networks are information processing systems. Neural networks can be thought of as "black box" devices that accept input and produce output. Each neural network has at least two physical components: connections and processing element (neuron). The combination of these two components creates neural networks. In a broad sense, neural networks consist of three principal elements:

Topology - how a neural network is organized into layers and how those layers are connected.

Learning- how information is stored in the network.

Recall - how the stored information is retrieved from the network.

In system identification view point, there are some advantages of neural networks to develop the model, (Sbarbaro, *et al.*1992).

Nonlinear system. Neural networks have greatest promise in the realm of nonlinear control problem. This stems from their ability to approximate arbitrary nonlinear mappings.

Learning and adaptation. Neural networks are trained using past data records from the system under study. Suitably trained networks have the ability to generalize when presented with inputs not appearing in the training data. Neural networks can also be adapted on line.

Multivariable systems. Neural networks naturally process many inputs and have many outputs. They are readily applicable to multi-variable process.

The neurons by themselves are not very powerful in terms of computation or representation but their interconnection allows encoding relations between the variables giving different powerful processing capabilities. The connection of several

layers gives the possibility of more complex nonlinear mapping between the inputs and the outputs. This capability can be used to implement classifiers or to represent complex nonlinear relations among the variables.

The most common of neural networks structure is Multi Layer Percepton (MLP). Figure 2 illustrates the example of MLP networks, which consist of input layer, hidden layer and output layer.

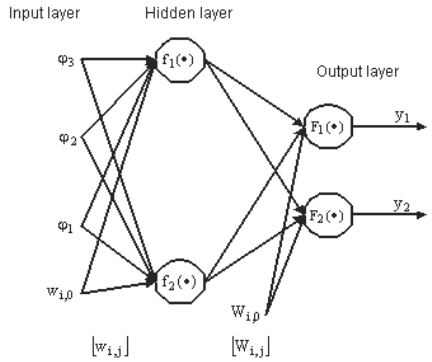


Figure 2. Structure of multilayer percepton

The mathematical formula expressing what is going on in the MLP networks takes the form:

$$y_i = F_i\left[\sum_{j=1}^{n_h} W_{i,j}.f_j\left(\sum_{l=1}^{n_\varphi} w_{j,l}\varphi_l + w_{j,0}\right) + W_{i,0}\right] \quad (13)$$

Cybenko (1989) shows that all continuous function can be approximated to any desired accuracy with neural networks of one hidden layer of hyperbolic tangent hidden neuron and a layer of linear output neuron.

In order to determine the weight values, a set of examples of how the outputs, $\hat{y}_i$, related to the inputs, φ_i, should be available. The task of determining the weights from these examples is called training or learning. The aim of training procedure is an adjustment of the weights to minimize the error between the neural networks output and the process output (also called target). A learning algorithm is associated with any change in the memory as represented by the weights; learning does not in this sense imply a change in the structure of the memory. Therefore, learning can be regarded as a parametric adaptation algorithm.

The learning algorithm that is used in this research is Levenberg Marquardt algorithm. Although, the Levenberg Marquardt algorithm is more complex than back propagation algorithm, but this algorithm can produce better results. Derivation of Levenberg Marquardt algorithm can be seen at Norgaard *et al.* (1999) and described as follows: Suppose that the training data are a set of input, u(k), and corresponding desired output, y(k). Specify the training set by:

$$Z^N = \{[u(k),y(k)] \mid k = 1, \ldots, N\} \quad (14)$$

The objective of learning is then to determine a mapping from the set of training data to the set of possible weights:

$$Z^N \rightarrow w$$

So that the networks will produce prediction, $\hat{y}(k)$, which in some sense are 'close' to the true output, y(k). The prediction error approach is based on the introduction of a measure of closeness in terms of a mean square error criterion:

$$V_N(w,Z^N) = L^{(i)}(w)$$
$$= \frac{1}{2N}\sum[y(k) - \hat{y}(k|w)]^T[y(k) - \hat{y}(k|w)] \quad (15)$$

The weights are then found as:

$$w = \arg\min_w V_N(w,Z^N) \quad (16)$$

By some kind of iterative minimization scheme:

$$w^{(i+1)} = w^{(i)} + \mu^{(i)}f^{(i)} \quad (17)$$

w^i specifies the current iterate, $f^{(i)}$ is the search direction and $\mu^{(i)}$ is the step size.

The Levenberg Marquardt is the standard method for minimization of mean square error criteria, due to its rapid convergence properties and

robustness. In this algorithm, there is a parameter, λ, to ensure convergence. The value of λ is controlled by the ratio between actual and predicted decrease:

$$r^{(i)} = \frac{V_N(w^{(i)}, Z^N) - V_N(w^{(i)} + f^{(i)}, Z^N)}{V_N(w^{(i)}, Z^N) - L^{(i)}(w^{(i)} + f^{(i)})} \qquad (18)$$

where:

$$L^{(i)}(w^{(i)} + f^{(i)}) = \sum_{k=1}^{N} \left(y(k)\hat{y}-(k/w) - f^T \frac{\partial \hat{y}(k/w)}{\partial w} \right)^2 = (\lambda^{(i)} f^{(i)T} f^{(i)}) - (f^{(i)T} G) \qquad (19)$$

G denotes the gradient of criterion with respect to the weights and R is the approximation of the Hessian. If the ratio is close to one, $L^{(i)}(w^{(i)} + f)$ is likely to be a reasonable approximation to V_N, and λ should be reduced by some factor. If, on other hand, the ratio is small or negative, λ should be increased.

The Levenberg Marquardt algorithm can be outlined as follows:

1. Select an initial weights vector $w^{(0)}$ and an initial value $\lambda^{(0)}$.
2. Determine the search direction from
$$\left[R(w^{(i)} + \lambda^{(i)} I \right] f^{(i)} = -G(w^{(i)})$$
3. $r^{(i)} > 0.75$ then $\lambda^{(i)} = \lambda^{(i)}/2$.
4. $r^{(i)} < 0.25$ then $\lambda^{(i)} = 2\ \lambda^{(i)}$.
5. If $V_N(w^{(i)} + f^{(i)}, Z^N) < V_N(w^{(i)}, Z^N)$ then accept $w^{(i+1)} = w^{(i)} + f^{(i)}$ as a new iterate and let $\lambda^{(i+1)} = \lambda^{(i)}$.
6. If the stopping criterion is not satisfied go to step 2.

Results and Discussion

The generating of training as well as validation data was carried out using Amplitude Pseudo Random Binary Signal (APRBS) as excitation signal in order to acquire the information of the process behavior for all amplitude and frequencies in its entire range of operation. Nelles and Isermann (1996) has shown that APRBS was the best excitation signal to emerge the information of process behavior. Due to the complexities of the process characteristics, special attention should be drawn to the parameters of APRBS, particularly in amplitude and pulse width of the signal, to cover the process characteristics. Figure 3 illustrates the input output data of the distillation column for training phase. The training data as well as the testing data were generated by applied the APRBS of input variables to the rigorous model of distillation column. The model was developed in Hysys software.

System identification of the distillation column was accomplished in NARX structure. Therefore, input variables of the model will consist of the present and past value of process input and output. The equation of model output can be expressed as follows:

$$\hat{Y} = f(Y_1, Y_2, U_1, U_2) \qquad (20)$$

where:

$$\hat{Y} = [\hat{y}_1(k+1)\ \hat{y}_2(k+1)]^T$$
$$Y_1 = [y_1(k), y_1(k-1), \ldots, y_1(k-ny_1)]$$
$$Y_2 = [y_2(k), y_2(k-1), \ldots, y_2(k-ny_2)]$$
$$U_1 = [u_1(k), u_1(k-1), \ldots, u_1(k-nu_1)]$$
$$U_2 = [u_2(k), u_2(k-1), \ldots, u_2(k-nu_2)]$$

ny and nu are the history length for output and input process, respectively.

The structure of neural networks is multi

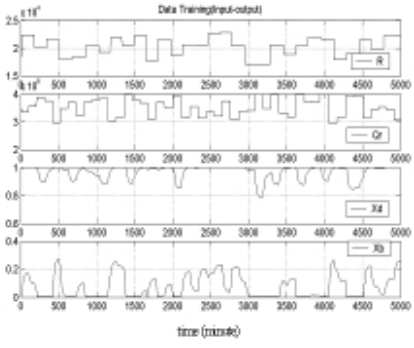


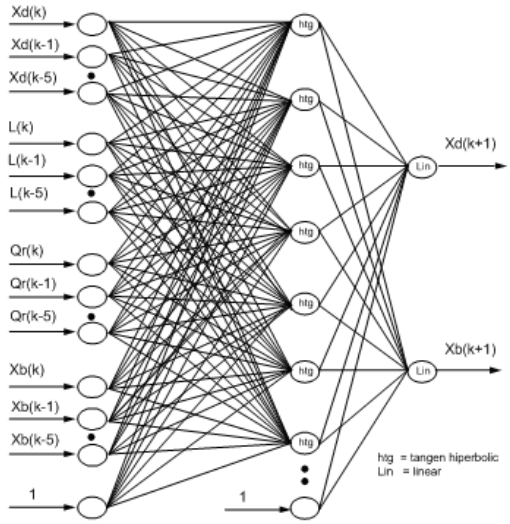
Figure 3. Data set for NN soft sensor training

layer percepton, which consists of three layers including input, hidden and output layer. Hyperbolic tangent function is selected as activation function in hidden layer, while the output layer is linear function. Some investigations were carried out to determine the history length and number of neurons in hidden layer. The best result was achieved when the number of neuron in hidden layer was 13 and value of history length for input and output was 6. Thus, there are 25 variables, in which the last variable is bias, for the input of neural model to produce two output variables. The structure of neural networks model is depicted in Figure 4.

The neural model should be trained with training data to determine fixed values of the weights. Then, the fixed weights will be used to validate the neural model using the other input-output data.

The goodness of system identification is measured using Root Mean Square Error (RMSE), which can be written as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}} \qquad (21)$$



Input layer (25 nodes) Hidden layer (13 nodes) Output Layer (2 nodes)

Figure 4. NN architecture

It is always desirable that the trained neural networks model is also validated on a set of data that was not used for training the networks. Using fixed values of the weights that obtained in training phase, the neural networks should produce the predicted output from the new input data. Figure 5 and 6 illustrate the output model for both output variables in validation phase.

Figure 5 is validation for the first output variable XD (mole fraction of distillate). In general, the neural networks model produces a good result. The neural networks model can still predict the dynamical behavior of the distillation column quite well. Thus, the model can identify and anticipate characteristics of the process. RMSE for the first output variable in validation phase is $6{,}6589 \times 10^{-5}$

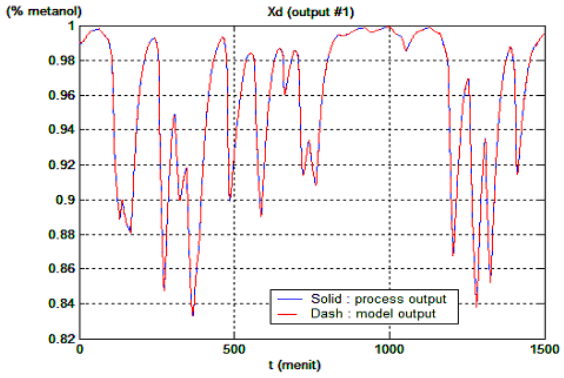


Figure 5 . Xd and NN output

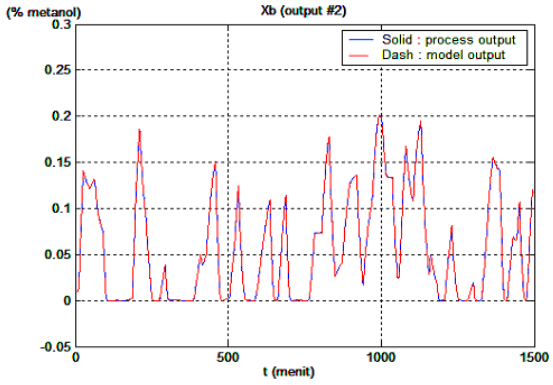


Figure 6. XB and NN output

In Figure 6, the neural networks model can also produce the good result to identify the second output variable X_B (mole fraction of bottom product) and yields RMSE = $1{,}98100 \times 10^{-4}$. Thus, the neural networks model can handle nonlinearities and complexities of the process.

Based on the above results, the neural networks model can handle the properties of distillation column. Moreover, it is also easy and simple to develop the nonlinear MIMO model using neural networks in the one model structure, so the neural network model can produce the predicted output faster than the analyzer, and requires less computational time.

Conclusion

The development of MIMO model using neural networks for distillation column, which consist of two input variables and two output variables, has been presented in this paper. The structure of the model is multilayer percepton.

In general, the designed neural networks can anticipate the nonlinearities and complexities of the distillation column. The neural networks model can produce the good results in modelling of distillation column with RMSE = $6{,}6589 \times 10^{-5}$ for the first output variable X_D (mole fraction of distillate) and RMSE = $1{,}98100 \times 10^{-4}$ for the second output variable X_B (mole fraction of bottom product). Therefore, the neural networks model can be used as an alternative model in developing a soft sensor.

NN soft sensor based on flow rate correlation is easy to build, has fast response, and no need as special instrumentations, has better reliability because flow rate measurement instrumentation has better reliability compared to analyzer reliability, is cheaper, and has low operational cost, low maintenance cost, and has good Root Mean Square Error (RMSE).

Reference

Biyanto, TR., Wahyudi, H., and Santosa, HH, 2004, Control Strategy of Single Distillation Column Binary System of Methanol-Water. Journal Instrumentasi, Instrumentation Society of Indonesia, Jakarta,

Biyanto, TR., and Santosa, HH, 2004, Modeling of methanol-water binary distillation column using a Neural Network, Journal Instrumentasi Vol 28 No 1, Instrumentation Society of Indonesia, Jakarta

Cybenko G, 1989, Approximation by Super-position of A Sigmoid Function, Mathematics of Control, Signal, and Systems, 2(4), 303-314

Kamohara, H, Takinami, MT, Takeda, M, Kano, M, Hasebe, S, and Hashimoto,I, 2003, Integration of Product Quality Estimation and Operating Condition Monitoring for Efficient Operation of Industrial Ethylene Fractionator, *IFAC*.

Luyben, W.L., 1990, Process Modeling, Simulation, and Control for Chemical Engineers, McGraw-Hill Inc. Singapore.

Luyben, W. L., Bjorn D. Tyreus, and Michael L. Luyben, 1998, *Plant wide Process Control*, Mc Graw-Hill, New York.

Nelles O, and Isermann R, 1996, Basis Function Networks for Interpolation of Local Linear Models, Proc.of 35 th Conference on Decision and Control, Kobe, Japan.

Norgaard, M,.Ravn, O., Poulsen, N.K., and Hansen L.K., 1999, Neural Network for Modelling and Control of Dynamic Systems, Springer London.

Ruiz J.P, Zumoffen D, Basualdo M, and Jimenez E., 2004, A Nonlinear Soft Sensor for Quality Estimation and Optimal Control Applied in a Ternary Batch Distillation Column, European Symposium on Computer Aided Process Engineering, Lisbon, Portugal.

Sbarbaro, D. Hunt, K.J, and Zbikowski, Rm, 1992, Neural Networks for Control Systems-A Survey, Automatica, 28(6), 1083-1112.

Zamprogna E., Barolo M. and Seborg D. E., 2000, Neural Network Approach to Composition Estimation in a Middle-Vessel Batch Distillation Column. *Proc. DINIP 2000.* Workshop on Nonlinear Dynamics and Control in Process Engineering, Rome.