Journal of Mining and Earth Sciences Vol.56 (2016) 84-92

Journal of Mining and Earth Sciences

Website: http://tapchi.humg.edu.vn

The application of artificial neural networks and backpropagation algorithm for flooding forecast models of petroleum distillation columns

Chi Van Dang[1,*], Phuong Mai Thi Phan[1]

[1]*Faculty of Electro-Mechanics, Hanoi University of Mining and Geology, Vietnam*

ARTICLE INFO

Article history:
Received 29 Jan. 2016
Accepted 02 Oct. 2016
Available online 20 Nov. 2016

Keywords:
Artificial neural network
Backpropagation algorithm
Distillation column

ABSTRACT

This paper presents the research and application of backpropagation algorithm of artificial neural network to create models that predict flooding of distillation column in the petroleum industry. The Matlab Simulink software has been used for modelling on the basis of the equations describing the process kinetics of the column, the survey of mathematical equations and evaluate the possibility of flooding. The data received from the modeling forms the basis for building neural network models. Results achieved in the modeling can be used for forecasting floods, as well as online measurement of flooding in distillation columns, which can assist the warning and prevent the incident occurred due to flooding of the column. As a result, the product quality, operation and safety of the distillation column in petroleum processing industry can be improved.

1. Introduction

In the past few decades, artificial neural network is a model based on the activities of the human brain, has been used effectively in the field of artificial intelligence, recognition, image processing, process signal, medicine and used widely in robot control, process control as well as in all fields of industry (Pham, 2009). In petroleum refineries and gas processing, distillation column is one of the most important processing stages that govern product quality and productivity of the plant. Important task of operating the column is the safety, product quality and prevention of the accidents that may occur. According to the assessment, flooding in the distillation column may be the most dangerous incident.

Flooding is overflow with fluid in the whole space between the discs due to the liquid that does not flow up and prevents gas from evaporation. Flooding makes distillation tower not only decommissioning but also may cause

*Corresponding author
E-mail: dangvanchi@humg.edu.vn

explosion of the tower due to the rising excessive pressure caused by the stasis accumulation of too many vapor phases. The column clogged in the event of flooding leads to the results in which the fluid in the flow line pipe can not flow up or because of the liquid on the surface of the disc is too violently under the action of the steam line up with the velocity great. So the study and survey of column inundation, application of artificial neural network to build predictive models is important in maintaining stable operation, warning, and preventing threats and incidents (Phan, 2002).

2. Artificial neural network and backpropagation algorithm

2.1. Artificial neural network

The works in (Dang, 2012) and (Phan, 2009) indicate that an artificial neural network (ANN) is a set of mathematical models similar to the structure of neurons in the brain. The ANN is formed from the linking of nerve cells called neurons and is arranged in a layer structure to form a network capable of performing parallel computation.

ANN can perform similarly to humans in terms of the reasoning, learning and storing the relationship of the process. ANN can draw an input and output space. Depending on the network structure, a series of connected neurons of the weights are adjusted to suit a variety of inputs to a wide range of outputs.

2.2. Backpropagation algorithm

The works (Dang, 2012), (Pham, 2009), and (Bryson and Ho, 1969) has proposed the usage of backpropagation (BP) learning rule for multilayer feedforward networks. The algorithm is used to train multi-layer network that have the processing elements of functions as nonlinear effects. Back propagation learning rule is the best training algorithm and is the popular choice today. Assume that input layer of the ANN includes m input neurons:

$$x = [x_1, x_2, \ldots, x_j, x_m]^T, \text{ with } j = 1, 2, \ldots, m$$

Output layer includes n output neurons

$$y = [y_1, y_2, \ldots, y_i, \ldots, y_n]^T, \text{ with } i = 1, 2, \ldots, n$$

The hidden layer can have one or more hidden layers, characterized neurons and are not in direct contact with input *x* and output *y*.

In the first phase: the sample input signal *x* spreads forward from input layer to hidden layer and then to output layer to form output signal *y*.

In the second phase: the deviation $e = (d-y)$ spreads in the opposite direction from the output layer to the hidden layers and then back into output layer, This phase has the task to adjust the weight values between the layers so that the output signal y follow output sample *d*.

In general, a straight transmission network has *S* layer; ${}^{s}net_i$ and ${}^{s}y_i$ is total weight and value of ithoutput neuron of sth layer; ${}^{s}w_{ij}$ is connection weights from ${}^{(s-1)}y_j$ to ${}^{s}y_i$ $(s = 1, 2, 3,\ldots,S)$ and has m input neuron layer and n output neuron layer. Backpropagation algorithm combines straight transmission network of s layer to find out the connection weights approximately adapted to the network and is described as follows:

Set the sample of in - out data: $[x^{(k)}, d^{(k)}]$ with $k = 1, 2, 3, \ldots, p$.

Step 1: Set the initial state of the network:

Select constant: $\eta > 0$;

Setting allowed deviation: E_{cp};

Setting the initial weights: $E = 0$ and $k = 1$;

Step 2: Cycle training:

Putting the k^{th} samples of input data in the input layer $(s = 1)$

${}^{s}y_i = {}^{1}y_i = x_i^{(k)}$ for all i.

Step 3: The straight transmission:

The process of signal transmission is presented a network via the Equation (1).

$$ {}^{s}y_i = a({}^{s}net_i) = a\left(\sum_j {}^{s}w_{ij}\, {}^{s-1}y_j\right) \qquad (1)$$

For each value *i* and *s* until the outputs ${}^{s}y_i$ of the output layer are determined.

Step 4: Calculate the output deviation:

Determining the value of deviations E_{new} and false signals ${}^{s}\delta_i$ for output neuron layer:

$$E_{new} = \frac{1}{2}\sum_{i=1}^{n}(d_i^{(k)} - {}^s y_i)^2 + E_{old} \quad (2)$$

where E_{new} and E_{old} are deviations in the new cycle and old cycle:

$${}^s\delta_i = (d_i^{(k)} - {}^s y_i)\, a'({}^s net_i) \quad (3)$$

Step 5: Backpropagation of deviation:

The process backpropagation of deviation is updated to calculate weights in terms of signal deviation ${}^{s-1}\delta_i$ for front layer:

$$\Delta^s w_{ij} = \eta\, {}^s\delta_i^{\,s-1} y_i \quad (4)$$

$${}^s w_{ij}^{new} = {}^s w_{ij}^{old} + \Delta^s w_{ij} \quad (5)$$

$${}^{s-1}\delta_i = a'({}^{s-1}net_i)\sum_j {}^s w_{ji}\, {}^s\delta_j \quad (6)$$

with s = S, S-1, ...,2.

Step 6: For each cycle of learning:

Check if $k < p$, then for $k = k + 1$ and return to step 1, otherwise to step 7.

Step 7: Check the total deviation.

If $E_{new} < E_{cp}$ then the training process finish and the results are recorded of the weight vector, deviation of the layers, and the required graph.

3. Development of mathematical equations and modeling of flooding tower

Distillation column is a nonlinear control object with multiple variables in the interaction with each other and the nature and evolution of the processes occur in the complex column. The column model will be surveyed to understand the characteristics and the phenomenons. The model is mainly based on theoretical models. In the paper, a deep study of the chemical and physical relationships is presented and the math equations are derived to describe the basic process. Modeling is provided by using the Matlab/Simulink software (Dang, 2012; Lutz, 2007).

Table 1. Operation factors of the C01 column

Number	Factors	Value	Number	Factors	Value
1	Height of tower	28m	4	Operating pressure	29 bar
2	Number of disc in theory	19 discs	5	Number of disc in fact	32 discs
3	Design temperature	-25°C →210°C	6	Temperature (top → bottom)	14°C → 98°C

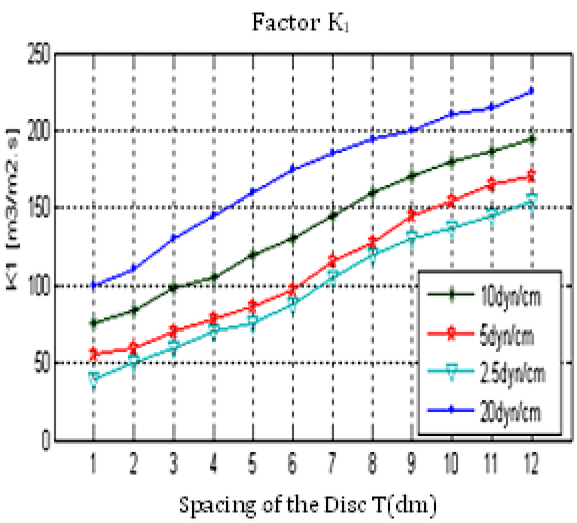


Figure 1. The pattern identification of the coefficient K_1

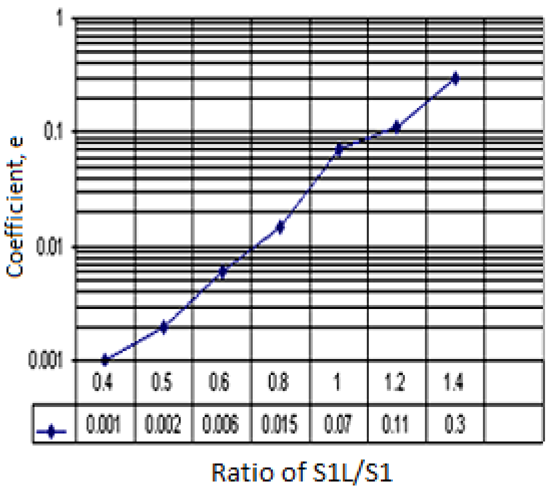
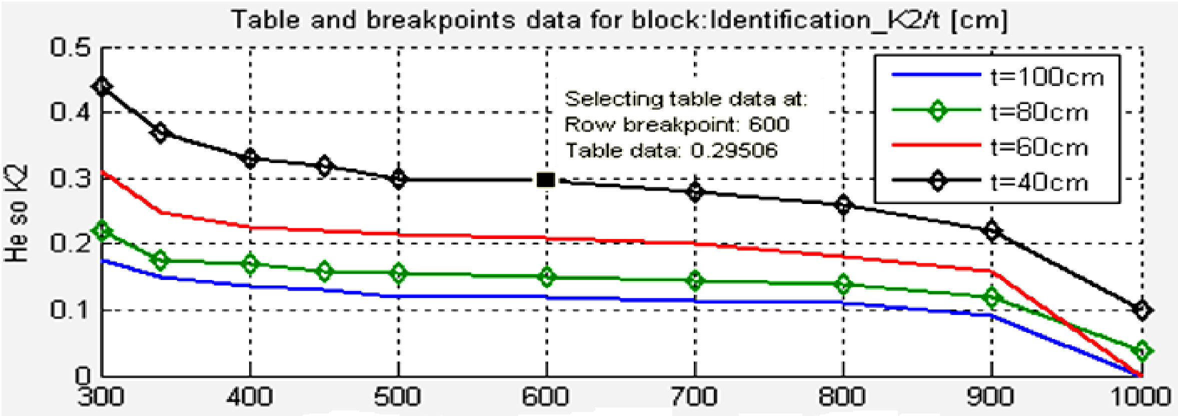


Figure 2. The pattern identification of the coefficient e



Specific weight liquid phase – vapor phase (kg/m³)

Figure 3. The pattern identification of the coefficient K_2

3.1. Development of mathematical equations

The author has studied and build math equations overview for a typical distillation column (Dang, 2012). The mathematical equations are generated in the steady state and the kinetic equation. However, the authors only introduce some mathematical equations to survey and evaluate the possibility of flooding column. Table 1 shows the specifications and operation of the C01 distillation column in Dinh Co Gas Processing Plant. This is ethane separation column – GPP working mode.

3.2. The mathematical equations for the flooding evaluation

For the evaluation of the possibility of flooding in the distillation column, a quantity is defined and called the non-flooded coefficient f. This variable is calculated by Equation (7).

$$f = \sqrt{\left(\frac{S_{1L}}{1.2S_1}\right)^2 + \left(\frac{S_{2L}}{1.8S_2}\right)^2} \quad (7)$$

where S_{1L}: The area of the capture area *(m^2); S_1:* The actual area in the region containing the capture *(m^2); S_{2L}:*The area of the flow tube *(m^2);S_2:* The actual area of the flow tube *(m^2).*

From the Equation (7), the flooding can not occur if the coefficient f satisfies the condition 0.7< f <1.2. However, it should not be satisfy the condition f < 0.7. Because the variables S_1 and S_2 that are too large will cause waste and the steam soar too slow. The term S_{1L} is calculated by Equation (8).

$$S_{1L} = \frac{Q_v}{K_1\sqrt{\left(\frac{\rho_L - \rho_V}{\rho_V}\right)}} \quad (8)$$

where K_1 depends on the distance between the disc and the surface tension, as shown in Figure 1 (Phan, 2002). The term S_{2L} is calculated by the following equation.

$$S_{2L} = K_2(Q_L + eS_1)$$

where *Q_v:* Steam flow on top of the column *(m^3/s), Q_L:* Liquid flow of theflowtube *(m^3/s)* *ρ_L, ρ_v:* Density of liquid flow, steam flow *(kg/m^3)*; e is iquid flow from 1m^2 of the disc surface in 1 minute, calculated by the m^3 and achieved from Figure 2 (Phan, 2002). The coefficient K_2 may be estimated by Figure 3, where t is distance between filter discs.

3.3. Flooding modeling and setting data for neural network.

3.3.1. Flooding modeling

The role of modeling will save time and reduce the number of experiments. In the distilling industry, the process of technology can not be observed. Therefore, the deep understanging of the core of the process is difficult. The modeling and simulation will help to understand the intrinsics of the process of technology. In the petrochemical refining industries, the software for simulation is often used such as ProII, Hysys. However, the existing softwares lack the tools to simulate the simulation identification, process diagnostics as the Fuzzy_Logic or Neuron_Network. For this reason, the Matlab/ Simulink software is used as a powerful tool in order to simulate the process and overcome the disadvantages.

In the paper, the verification of the model is implemented. Then, the comparision between output results from the model and the data from the C01 distillation column of Dinh Co Gas Processing Plant is executed. The interface of simulation and inundation survey column are described In Figure 4.

The influence of steam flow to the possibility of flooding tower is shown in Figure5

3.3.2. Setting data for neural network

The modeling and assessment surveys flooding multi-component distillation column C01 is performed. Datasets measured is recorded in *Workspace in Matlab*, which makes database to train the neural network (Pham, 2009) and (Lutz, 2007). They are divided into two sets, one for training network and one for checking predictive models. Flooding distillation column was assessed by the coefficient of flooding in the Equation (7), including the geometry parameters of column, steam flow on the top of column, fluid flow in the column, the density of liquid flow, steam flow ... But research content is limited to a few fixed operating conditions that are only survey of influences of steam flow, liquid flow in pipes with the distance between the filter disc T = 80cm .

The dependence of steam flow on the top of column Q_V, liquid flow in the flow pipes Q_L to pass coefficient column flooding is shown in Figure 6

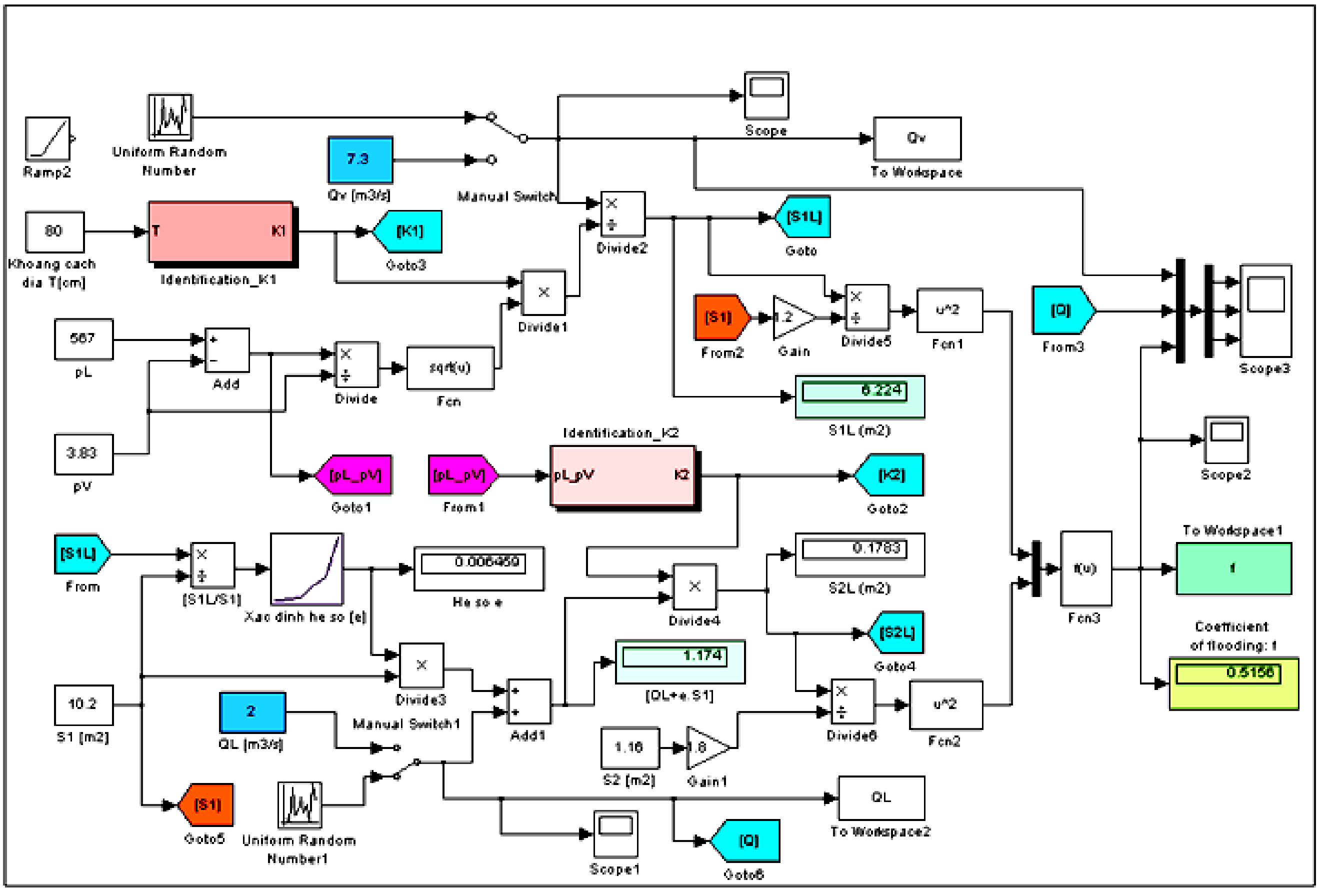


Figure 4. The interface Matlab Simulink survey column inundation C01

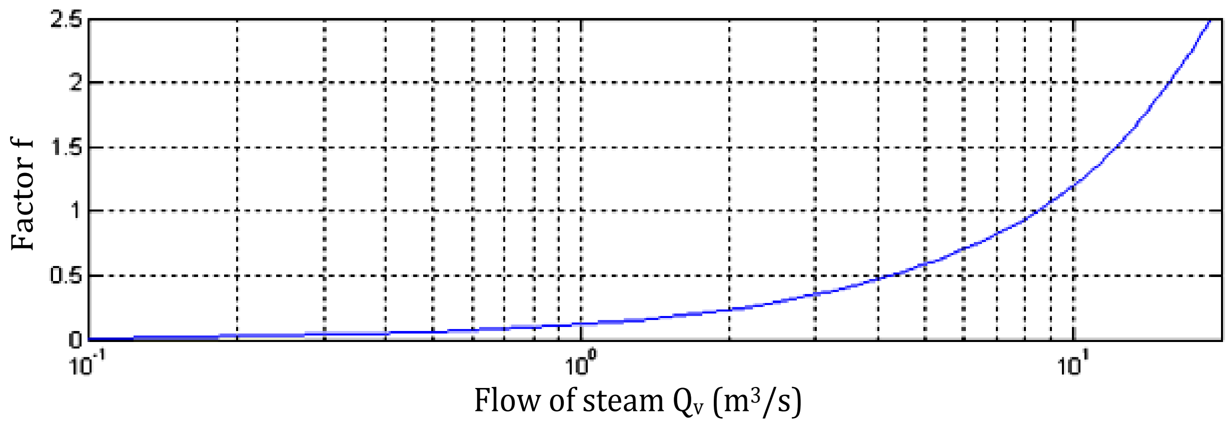


Figure 5. The influence of steam flow to the possibility of the flooding column

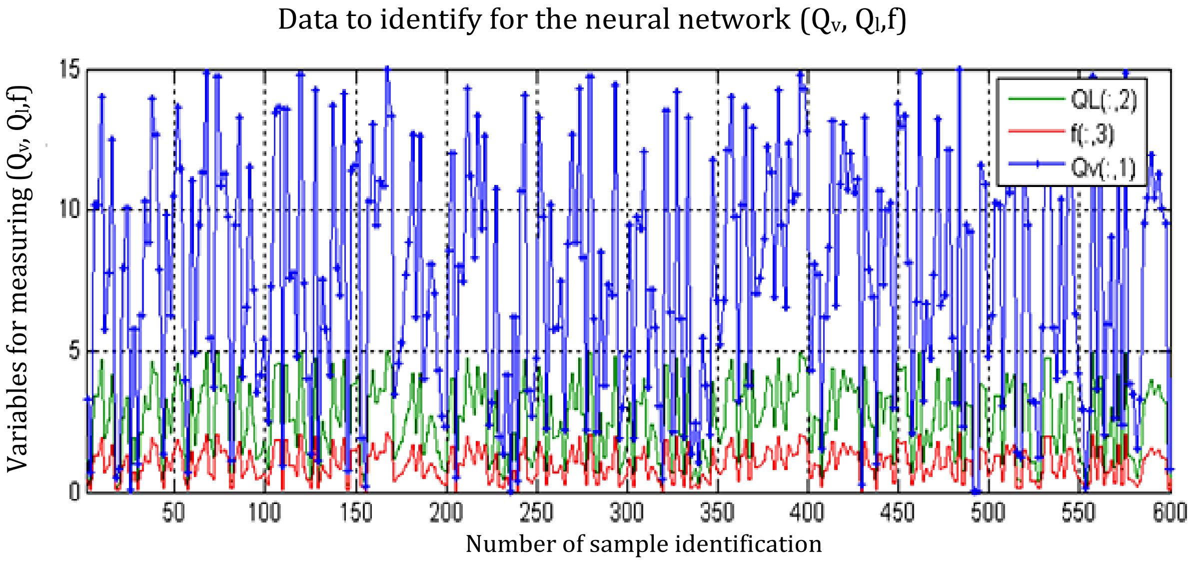


Figure 6. The influence of Q_v, Q_L to flooding f

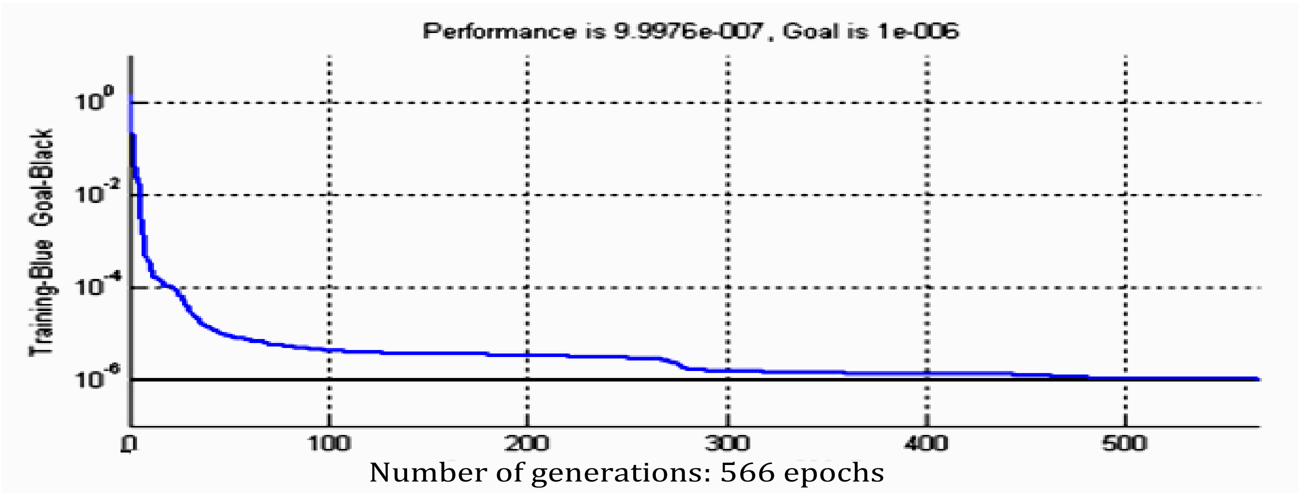


Figure 7. The deviation of training process

4. Application of neural network to build predictive models

4.1. Training and selecting configuration neural network

The training process of network is performed by the Matlab software with many different network structures. In this application, the training process is selected via a configuration that is suitable for the dataset. The training implementation of a feedforward neural network has 3 layers, which is written by code in MATLAB m-files (Samsung) as follows:

```
% Training program for neural network
inp=P'; %training Patterns
out=T'; %training Targets
net=newff(minmax(p),[3,7,1],{'tansig','tan
sig','purelin'},'trainbfg');
net.trainParam.show = 10;
net.trainParam.epochs = 1000;
net.trainParam.lr = 0.05;
net.trainParam.goal = 1e-6;
[net,tr]=train(net,P,T);
net=train(net,inp,out);
gensim(net,0.001);
…..
```

In which:

- Input layer has 3 neurons of the use of tansig transfer function;
- Hidden layer has 7 neurons of the use of tansig transfer function;
- Output layer has 1 neuron of the use of purelin transfer function;
- The value of bias in the process of learning E=1.e^{-6};
- Learning rule: backpropagation algorithm;
- The number of generation for training of 566 epochs

4.2. The result of modeling

- The influence of the epochs trained to bias the pattern in the training process is shown in Figure 7 of the network structure of 7 neural hidden layers (Figure 8).
- The result forecasts of neural network model are shown Figure 9.

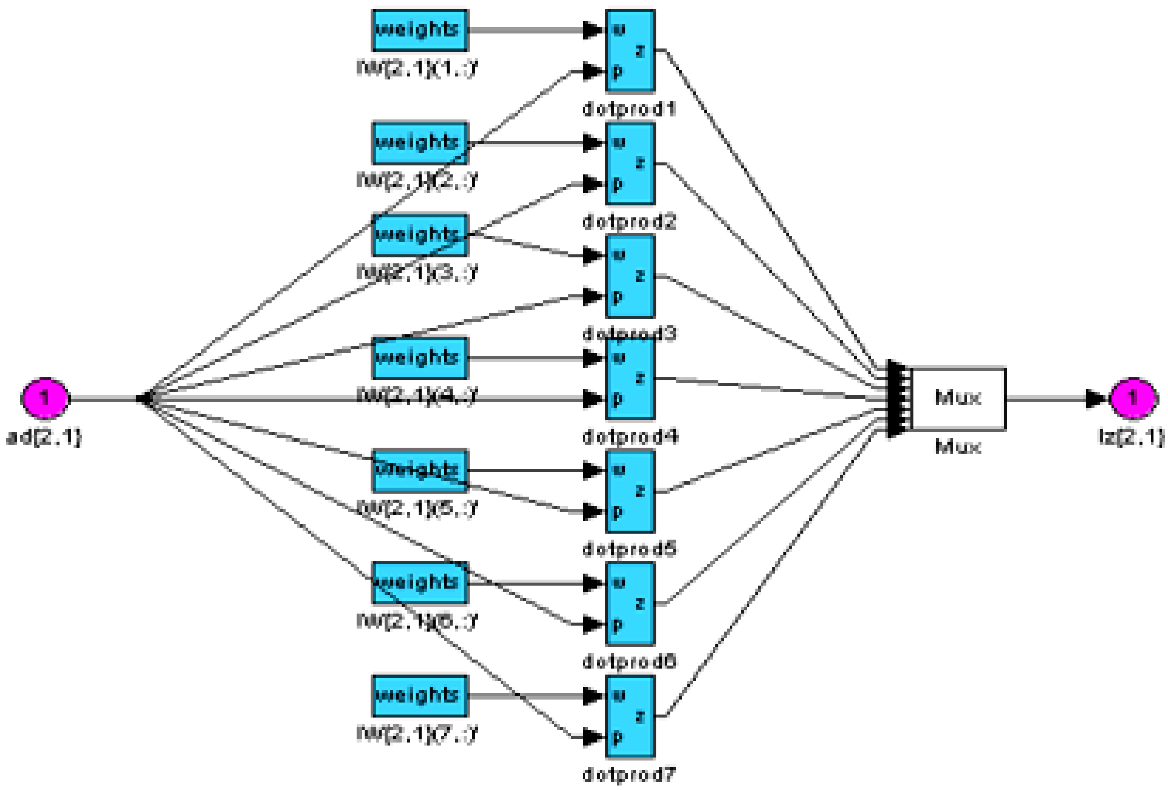


Figure 8. The structure neuron network of hidden layer

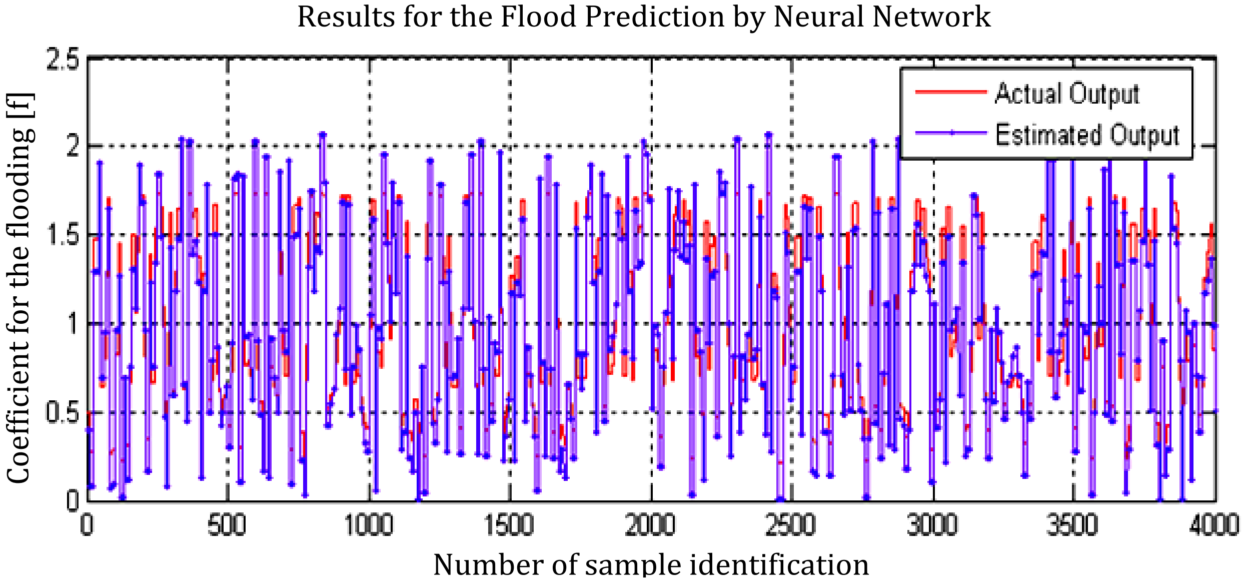


Figure 9. Results of the neural network forecast model

5. Conclusion

With the actual operating parameters of the C01 column, the coefficient of flooding f = 0.872 is smaller than 1.2 and greater than 0.7. This coefficient ensures for a normal operation of column that is not flooded.

When changing the steam flow, the coefficient of flooding is changed according to characteristics, as shown in Figure 6. At this time, the range can be adjusted the condition $Q_{v(max)}<10m^3/s$ such that column is not flooded.

Deviations between the predicted value and the actual value are achieved of a setting of training e = 10^{-6} after 566 epochs of training (see Figure 8).

Figure 9 shows the results predicted by the network model with the in - out data set of 3-layer neural network model. The model (3x7x1) are adhesive according to the samples of in - out signals. Thus, an feedforward neural network of 3 layers are confirmed for learning the samples in - out signals.

The backpropagation learning rule in neural network was built and trained according to set objectives. After the training, it will be installed into the library: Simulink\Neural Network blockset\Neural Network. ANN can be installed and run parallelly with simulink models of distillation column C01. Modeling can survey and evaluate influence on operating parameters to coefficient of flood column, and help operators to adjust parameters that are suitable for work to ensure the safety in working and efficiency, and also prevent incidents of flooding that may occur.

Predictive models only address the influence of steam flow column, liquid flow in pipe flow to pass and disk distance to coefficient of flooded column. Other parameters will be further studied by the complex relationship between steam flow to the feed flow, pressure and temperature column, reflux ratio of top and bottom flow column.

References

Dang, V.C., 2012. *Study on methods to improve product's quality and efficiency of the distillation column in petroleum industry*. Unpublished Engineering Doctoral Thesis, Hanoi University of Mining and Geology, Hanoi.

Lutz, W., 2007. *Taschenbuch der Regelungstechnik: Mit MATLAB und Simulink*. Harri Deutsch, Deutsch.

MathWorks, 2005. *Neural Network Toolbox*. Help in MATLAB, © 1994-2005.

Pham, H.D., 2009. *Neural network and application in automatic control.* Publisher of Scientific and Technical, Hanoi.

Phan, T.B., 2002. *Technology textbooks for oil refinery*. Publisher of Building, Hanoi, 18-47.

Ramesh, K., Aziz, N. And Shukor, S.R., 2007. Dynamic Rate-Based and Equilibrium Model Approaches for Continuous Tray Distillation Column. *Journal of Applied Sciences Research* 3(12), 2030-2041.

Samsung Engineering Condensate, 1987. *Guide the operation of the Dinh_Co Gas Processing Plant* Ltd_Operating. Manual Rev_D_1987.