IJCPE
Iraqi Journal of Chemical and Petroleum Engineering

Iraqi Journal of Chemical and Petroleum Engineering
Vol.10 No.4 (December 2009) 57-62
ISSN: 1997-4884

University of Baghdad
College of Engineering

Prediction of the Point Efficiency of Sieve Tray Using Artificial Neural Network

Prof. Dr. Adil. A. Al-Hemiri* and Firas N. Hassan

**Chemical Engineering Department - College of Engineering - University of Baghdad – Iraq*

Abstract

An application of neural network technique was introduced in modeling the point efficiency of sieve tray, based on a data bank of around 331 data points collected from the open literature. Two models proposed, using back-propagation algorithm, the first model network consists: volumetric liquid flow rate (QL), F factor for gas (FS), liquid density (ρL), gas density (ρg), liquid viscosity (μL), gas viscosity (μg), hole diameter (dH), weir height (hw), pressure (P) and surface tension between liquid phase and gas phase (σ). In the second network, there are six parameters as dimensionless group: Flow factor (F), Reynolds number for liquid (ReL), Reynolds number for gas through hole (Reg), ratio of weir height to hole diameter (hw/dH), ratio of pressure of process to atmosphere pressure (P/Pa), Weber number (We). Statistical analysis showed that the proposed models have an average absolute relative error (AARE) of 9.3% and standard deviation (SD) of 9.7% for first model, AARE of 9.35% and SD of 10.5% for second model and AARE of 9.8% and SD of 7.5% for the third model.

Keywords: Sieve tray, Point efficiency, artificial neural network, Back-propagation algorithm

Introduction

Perforated trays have been used for many years for liquid –gas contacting in commercial distillation or absorption columns. Perforated trays came in widespread use, in the 1950-1960 period, because columns provided with them were shown to have (1) :

- Higher throughput compared with those provided with bubble-cap trays.
- Ability to handle suspended solid particles (with suitable hole size), flushing them down from tray to tray.
- Compactness especially in cryogenic applications.
- Satisfactory operation range, typically from 50-120 percent of design capacity.
- Low pressure drop, so they are preferable for vacuum applications. Maximum pressure drop for vacuum distillation usually ranges from 4 to 8 cm of liquid.
- Lower cost in comparison with bubble-cup tray. The relative cost will depend on the material of construction used; for mild steel the ratios bubble-cap : valve : sieve are 3.0 : 1.5 : 1.0.

For these reasons, sieve trays are cheapest and satisfactory for most applications. However, valve trays should be considered if the specified turndown ratio can not be met with sieve trays. Bubble-cap trays should be considered when very low gas rates have to be handled and positive liquid seal is essential at all flow rates.

Tray efficiency

The designer is concerned with real contacting stages; not the theoretical equilibrium stage assumed for convenience in the mathematical analysis of multistage processes.

Equilibrium will rarely be attained in a real stage. The concept of a stage efficiency is used to link the performance of practical contacting stages to the theoretical equilibrium stage[(1)].

Three principal definitions of efficiency are used:

1. Murphree plate efficiency, defined in terms of the vapour compositions by:

$$E_{mV} = \frac{y_n - y_{n-1}}{y_e - y_{n-1}}$$

...... (1)

where *ye* is the composition of the vapour that would be in equilibrium with the liquid leaving the plate. The Murphree plate efficiency is the ratio of the actual separation achieved to that which would be achieved in an equilibrium stage (see fig. 1) . In this definition of efficiency the liquid and the vapour stream are taken to be perfectly mixed; the compositions in equation (1) are the average composition values for the streams.

2. Point efficiency (Murphree point efficiency). If the vapor and liquid compositions are taken at a point on the plate, equation 1 gives the local or point efficiency,(E_p).

3. Overall column efficiency. This is sometimes confusingly referred to as the overall Plate efficiency.

$$E_0 = \frac{\text{Number of ideal stage}}{\text{Number of real stage}} \qquad (2)$$

An estimate of the overall column efficiency will be needed when the design method used gives an estimate of the number of ideal stages required for the separation.

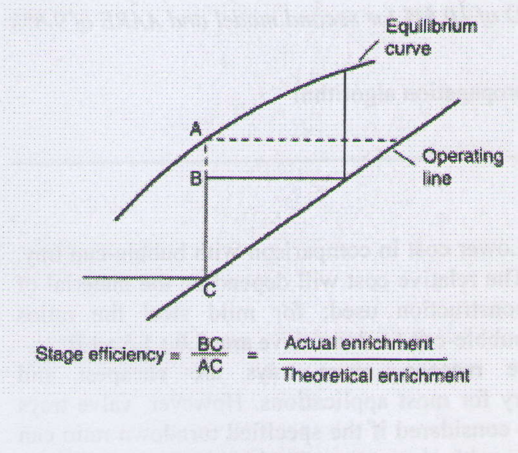
Equilibrium curve
A
B
C
Operating line
Stage efficiency = $\frac{BC}{AC}$ = $\frac{\text{Actual enrichment}}{\text{Theoretical enrichment}}$

Fig. 1 Plate efficiency

In some methods, the Murphree plate efficiencies can be incorporated into the procedure for calculating the number of stages and the number of real stages determined directly.

For the idealized situation where the operating and equilibrium lines are straight, the over all column efficiency and the Murphree plate efficiency are related by an equation derived by Lewis (1936):

$$E_0 = \frac{\log[1 + E_{mV}(\frac{mV}{L} - 1)]}{\log(\frac{mV}{L})} \qquad (3)$$

Where m = slope of the equilibrium line,
V = molar flow rate of the vapor,
L = molar flow rate of the liquid.

Feedforward, Back-Propagation

The feed forward, back-propagation architecture was developed in the early 1970's by several independent sources. This independent co-development was the result of a proliferation of articles and talks at various conferences which stimulated the entire industry. Currently, this synergistically developed back-propagation architecture is the most popular, effective, and easy to earn model for complex, multi-layered networks. This network is used more than all other combined. It is used in many different types of applications. This architecture has spawned a large class of network types with many different topologies and training methods. Its greatest strength is in non-linear solutions to ill-defined problems.

The typical back-propagation network has an input layer, an output layer, and at least one hidden layer. There is no theoretical limit on the number of hidden layers but typically there is just one or two. Some work has been done which indicates that a minimum of four layers (three hidden layers plus an output layer) are required to solve problems of any complexity. Each layer is fully connected to the succeeding layer, show in Fig. (2).

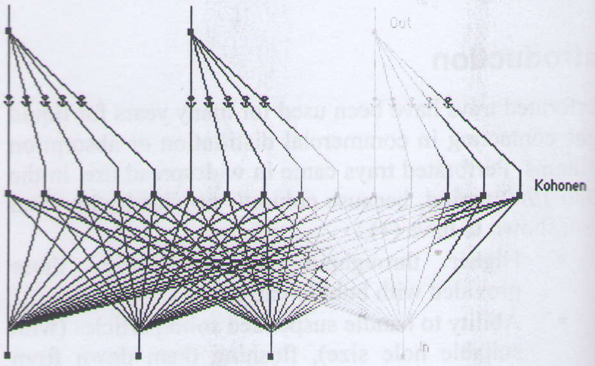

Fig. 2An Example Feed forward Back-propagation Network

Training a Back-Propagation Network

The conventional algorithm used for training a MLFF is the Bp algorithm, which is an iterative gradient algorithm designed to minimize the mean-squared error between the desired output and the actual output for a particular input to the network [2].

Two learning factors that significantly affect convergence speed as well as accomplish avoiding local minima, are the learning rate and momentum.

The learning rate (η) determines the portion of weight needed to be adjusted. However, the optimum value of η depends on the problem. Even though as small learning rate guarantees a true gradient descent, it slows down the network convergence process. If the chosen value of η is too large for the error value, the search path will oscillate about the ideal path and converges more slowly than a direct descent. The momentum (α) determines the fraction of the previous weight adjustment that is added to current weight adjustment. It accelerates the network convergence process. During the training process, the learning rate and the momentum bring the network out of its local minima, and accelerate the convergence of the network.The algorithm of the error back-propagation training is given in the Appendix

Computer Simulation Results

Collection of Data

The first step in neural network modeling is the collection of data. The data are necessary to train the network and to estimate its ability to generalize. In this model about 331 experimental points have been collected to predict sieve tray efficiency. The data were divided into training and test sets: the neural network was trained on 85% (280 points) of the data and tested on 15% (51 points). Table 1 gives the detailed listing of data used for the present work.

The ANN structure is determined by trial and error. For the first model it consists of ten input neurons in the input layer, twenty three neurons in the hidden layer and one neuron in the output layer, for second case it consists of seven neurons in the input layer, thirteen neurons in the hidden layer and one neuron in the output layer. Then the networks were trained with back-propagation algorithm and the weights and biases matrices were calculated. The best ANN models for the two cases are shown in tables (2) and (3).

Table 1 Details of data bank used for the present work.

No. of data	System	Author
33	i butane/n butane at 2758 Kpa	Sakata &Yangi[3] (1979)
35	i butane/n butane at 2068Kpa	Sakata &Yangi[3] (1979)
17	i butane/n butane at 1138Kpa	Garcia & Fair (2000)
21	i butane/n butane at 1138Kpa	Yangi &Sakata[4] (1982)
19	Cyclohexane/n heptane at 165Kpa	Garcia & Fair[5] (2000)
22	Cyclohexane/n heptane at 165Kpa	Yangi & Sakata[4] (1982)
8	Cyclohexane/n heptane at 414Kpa	Garcia & Fair[5] (2000)
18	Cyclohexane/n heptane at 165Kpa	Sakata &Yangi[3] (1979)
18	Air/water at 101.4Kpa	Nutter[6] (1972)
15	Cyclohexane/n heptane at Kpa	Yangi & Sakata[4] (1982)
18	Ortho/paraxylenes at 2.13Kpa	Bennett & Agrawal[7] (1983)
18	Methyl isobutyl keton/water at 101.4Kpa	Rush & Stirba[8] (1957)
13	2-propanol/water 13.3Kpa	Bennett & Agrawal[7] (1983)
10	Ethylene glycol/water at 6.7Kpa	Garcia & Fair[5] (2000)
11	n-octanol/n-decanol at 1.3Kpa	Bennett & Agrawal[7] (1983)
23	Air/water at 101.4Kpa	Harada & Adachi[9] (1964)
32	Air/water at 101.4Kpa	Muhanad.A.R.Hassan[10] (2005)

Table 2 finding the best ANN model for case 1

Structure	MSE	No.of iteration	Learning rate	Momentum coefficient	Transfer function
[10-23-1]	0.001	2497	0.8	0.9	tan sigmoid

Table 3 finding the best ANN model for case 2

Structure	MSE	No. of iteration	Learning rate	Momentum coefficient	Transfer function
[6-30-1]	0.002	11423	0.75	0.9	tan sigmoid

Test of the Proposed ANNs

The ANN models were tested using another set of data to show the accuracy of the network for predicting point efficiency of sieve tray. The two models were used to generate (51) new data values each, but in the case of the second model, the data values were as dimensionless groups.

The comparison between experimental and predicted efficiency for the two cases are plotted in Figures (3) and (4).

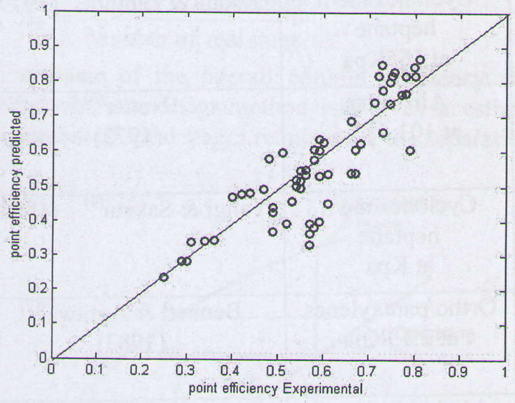

Fig. (3) Comparison between experimental and prediction point efficiency for case (1) in testing set.

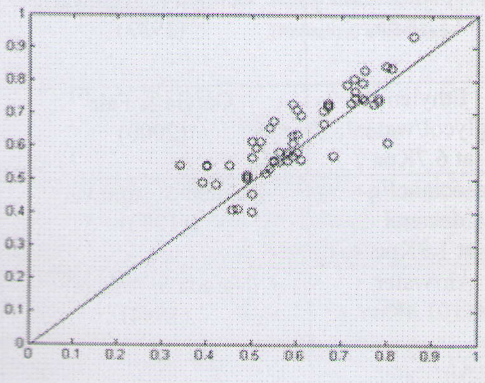

Fig. (4) Comparison between experimental and prediction Point efficiency for case 2 in testing set.

Discussion

This work used high performance algorithm presented by back-propagation, which uses the (trainlm) function, trainlm can train any network as long as its weight, net input, and transfer functions have derivative functions. Trainlm is the default training function because it is very fast, but it requires a lot of memory to run. Training stops when any of these conditions occur:

- The maximum number of epochs (repetitions) is reached.
- The maximum amount of time has been exceeded.
- Performance has been minimized to the goal.
- The performance gradient falls below mingrad.

The learning rate is constant throughout training, if it is set too high, the algorithm may oscillate and become unstable. But if it is set too small, the algorithm will take too long to converge. The learning rate could be changed if needed, and in our study a learning rate of 0.8 for case 1 , 0.75 for case 2 and 0.65 for case 3. The selection of the number of hidden layer neuron is very important and troublesome. If the number of the hidden layer neuron is fewer, the ANN cannot receive all necessary information of the modeling system and has less tolerance on faults, so that it gives wrong output. On the contrary, the ANN may cause a phenomenon called "over fitting". In this study, the first network consisted of 23 neuron in the hidden layer and the targeted error (MSE=0.001) was achieved, while the second network consisted of 30 neurons but gave, somehow, slightly larger error (MSE=0.002). However both correlations represented the data quite well as can be seen from the values of the standard deviation (SD) and the average absolute relative error (AARE).

Conclusions

It has been demonstrated that the optimal model is a network with one hidden layer. The application of ANN to the sieve tray point efficiency indicated the coming of a flexible tool for engineers.

A neural network model has been developed for the prediction of the sieve-tray point efficiency. The purely empirical model was tested on data that were not used to train the neural network and yielded very accurate predictions for the two cases considered.

- Case 1: the model, $E_p \propto F_s, Q_L, \rho_g, \rho_L, \mu_g, \mu_L, \sigma, h_w, d_H, P$ has a structure of 10-23-1 with AARE of 9.3%.
- Case 2: the model, $E_p \propto F, Re_L, Re_g, We, (h_w/d_H), (P/P_a)$ has a structure of 6-30-1 with AARE of 9.35%.

Nomenclature

b = Bias.

d= $_{Hole}$ diameter.

E_{MV}= Murphree gas-phase tray efficiency.

$E_{P=}$ Point efficiency.
E_0 = Overall column efficiency.

f = The activation function.

f' = The derivation of the activation function.

F = Flow factor= $(V_L/V_g)(\rho_L/\rho_g)^{0.5}$.

F_s = F factor base on superficial velocity of gas = $V_g(\rho_g)^{0.5}$, $(cm/s)(g/cm^3)^{0.5.}$

h_j = The actual output of hidden neuron j.

h_w= Outlet weir height , cm.
L = Liquid flow rate, mole/s.
m = Slope of equilibrium line.

o_k = The actual output of output neuron k.

Q_g = Volumetric gas flow rate, cm^3/s.
Q_L = Volumetric liquid flow rate, cm^3/s.
Re_g = Reynolds number for gas =$(\rho_g u_g d_H/\mu_g)$.
Re_L = Reynolds number for liquid =$(\rho_L u_L d/\mu_L)$
V =Gas flow rate, mole/s.

W_{ij} = Synaptic weights between input and hidden Neurons.

W_{jk} = Synaptic weight between hidden and output neuron.

We = Weber number=$(\Delta\rho V_{gh}^2 d_h/\sigma)$.
x = Liquid concentration (mole fraction).
y = Gas concentration (mole fraction).
y^*_n = Mole fraction that would be in the equilibrium with liquid leaving tray.
$\overline{y_n}$= Average mole fraction in the vapor leaving the tray.

Greek Symbols

δ_k = The error term.
λ = Stripping factor =(m V/L).
μ_g =Gas viscosity, g/cm.s .
μ_L= Liquid viscosity , g/cm.s .
ρ_g = Gas density , g/cm^3.
ρ_L = Liquid density ,g/cm^3.
σ = Surface tension, g/s^2.

Abbreviations

AARE = Average Absolute Relative Error.
ANN = Artificial Neural Network.
BP = Back Propagation.
MLFF = Multi-Layer Feed Forward.
MSE = Mean Square Error.
S.D = Standard Deviation.

References

1- Coulson, J.M.; Richardson, J.F. and Sinnott, R.K. (1999), "Chemical Engineering" Vol.6, 3rd ed.,.

2- Lendaris, G. (2004), “Supervised learning in ANN from introduction to artificial intelligence”, New York, April 7.

3- Sakata, M. and Yanagi, T. Performance of a Commercial Scale Sieve Tray. Inst. Chem. Eng. Symp. Ser. 1979, 56, 3.2/21.

4- Yanagi, T. and Sakata, M. Performance of a Commercial Scale 14% Hole Area Sieve Tray. Ind. Eng. Chem. Process Des. Dev. 1982, 21, 712.

5- Garcia, J. and Fair, J. R. (2000), "A fundamental model for prediction of distillation sieve tray efficiency. Model development and validation", Ind. Eng. Chem. Res., 39, 1818.

6- Nutter, D. E. Ammonia Stripping Efficiency Studies. AIChE. Symp. Ser. 1972, 124, 68.

7- Bennett, D. L.; Agrawal, R. and Cook, P.J.(1983), "New pressure drop correlation for sieve tra distillation columns" AIChEJ, 29, 3 434-442

8- Rush, F. and E.; Stirba C. Measured Plate Efficiencies and Values Predicted from Single-Phase Studies. AIChE J. 1957, 3, 336.

9- Harada, M. and M. Adachi, Studies of fluid mixing on sieve plate. International Chemical Engineering. 4, No1

10- Muhanad .A.R.Hassan ,(2005)"Prediction of Point Efficiency of sieve tray Using Air Water System , Ph.D. Thesis,University of Baghdad.

Appedix

Step1: initialize network weight values.

Step2: sum weighted input and apply activation function to compute output of hidden layer.

$$h_j = f\left[\sum_i X_i W_{ij}\right] \qquad (i)$$

Where

h_j: The actual output of hidden neuron j for input signals X.

X_i: Input signal of input neuron (i).

W_{ij}: Synaptic weights between input neuron hidden neuron j and i.

f: The activation function.

Step3: sum weighted output of hidden layer and apply activation function to compute output of output layer.

$$O_k = f\left[\sum_j h_j W_{jk}\right] \qquad (ii)$$

Where

O_k: The actual output of output neuron k.

W_{jk}: Synaptic weight between hidden neuron j and output neuron k.

Step4: compute back propagation error.

$$\delta_k = \left(d_k - O_k\right) f'\left(\sum_j h_j W_{jk}\right) \qquad \text{(iii)}$$

where

f' :The derivative of the activation function.

d_k: The desired of output neuron k.

Step5: calculate weight correction term.

$$\Delta W_{jk}(n) = \eta \delta_k h_j + \alpha \Delta W_{jk}(n-1) \qquad \text{(iv)}$$

Step6: sums delta input for each hidden unit and calculate error term.

$$\delta_j = \sum_k \delta_k W_{jk} f'\left(\sum_i X_i W_{ij}\right) \qquad \text{(v)}$$

Step7: calculate weight correction term.

$$\Delta W_{ij}(n) = \eta \delta_j X_i + \alpha \Delta W_{ij}(n-1) \qquad \text{(vi)}$$

Step8: update weights.

$$W_{jk}(n+1) = W_{jk}(n) + \Delta W_{jk}(n) \qquad \text{(vii)}$$

$$W_{ij}(n+1) = W_{ij}(n) + \Delta W_{ij}(n) \qquad \text{(viii)}$$

Step9: repeat step2 for given number of error.

$$MSE = \frac{1}{2p}\left[\sum_p \sum_k \left(d_k^p - O_k^p\right)^2\right] \qquad \text{(ix)}$$

Where p: The number of patterns in the training set.

Step10: END