Packed Column Simulation using Various Advanced Techniques

A. Shahsavand, F. Derakhshanfard, F. Setoudeh
Department of Chemical Engineering, Faculty of Engineering
Ferdowsi University of Mashhad
Mashhad, Iran

Abstract

Various sets of experimental data collected from a pilot scale packed absorption column are used to compare the generalization performances of the Back Propagation Multi-Layer Perceptron (BPMLP) and the Radial Basis Function (RBF) neural networks. The 11cm diameter packed tower filled with 1.8 meter ¼ inch ceramic Rashig rings was used for separation of carbon dioxide from air using various concentrations and flow rates of Di-Ethanol Amine (DEA) and Methyl Di-Ethanol Amine (MDEA) solutions. Two efficient algorithms were employed for optimal training of both neural networks. The Leave One Out Cross Validation (LOOCV) criterion was employed to compute the optimum level of regularization for RBF networks. An in-house procedure was also exploited to predict the optimal widths of isotropic Gaussian basis functions for these networks. Another in-house algorithm was used to train the MLP networks more rapidly and efficiently than the conventional procedures. The simulation results indicated that the RBF networks can perform more adequately than the MLP networks. Because, the RBF networks enjoy a solid theoretical background which enables them to successfully filter out the noise and provide more reliable generalization performances.

Keywords: Absorption, Unit operations, Optimization, Packed bed, RBF, MLP

1. Introduction

Separation of carbon dioxide from air and various industrial gases is essential from both operational and environmental views. For example, CO_2 must be separated from natural gas to increase its heating value or carbon dioxide is usually extracted from various flue gases in beverage industry. To reduce global warming, CO2 should be also removed from industrial flue gases before exhausting them to atmosphere.

Although several adsorption and membrane processes are recently used for CO2 separation purposes [1-5] absorption processes are still more popular in this area [6-8] Alkanolamines (such as DEA or MDEA) are usually used for efficient separation of carbon dioxide from various industrial gases Packed towers are usually provide higher mass transfer areas and lower pressure drops when compared to tray towers.

Lin and Shyu [9] investigated the absorption of carbon dioxide from nitrogen using MEA and MDEA solutions in a packed column under various operating conditions. A two parameter theoretical model was presented for describing the CO_2 absorption behavior. The proposed model was validated using test data.

Sultan et al [10] presented a theoretical model to investigate the effect of various operational parameters on the performance of a regeneration packed column. The experimental data were then correlated to estimate the water evaporation rate from desiccant ($CaCl_2$) at various operating conditions. They concluded that "the water evaporation rate increases with increase of air and solution inlet parameters, namely, flow rate and temperature".

Sharma et al [11] employed back propagation artificial neural networks to investigate the fault diagnosis in an ammonia–water packed distillation column. The network was reported to perform satisfactorily on detection of the designated faults. The relative importance of various input variables on the output parameters was calculated by partitioning the connecting weights. The simulation results indicated that "bottoms temperature, overhead composition and overhead temperature are not much affected by the disturbances in feed rate, feed composition and vapor rate in the given range".

2. Experimental Data

After calibration of the gas analyzer apparatus, various measurements of inlet and outlet gas concentrations were performed. The following operating variables were varied during the experiments:

- ✓ Type and solvent (DEA, MDEA and pure water),
- ✓ Gas and liquid flow rates,
- ✓ Concentrations of both solvents and gas streams.

The small temperature fluctuations during each experiment were ignored and the average temperature was calculated using the initial and final conditions. The barometric pressure was close to one standard atmosphere for all experiments. Figures 1 and 2 represent typical collected data for the operating conditions specified in the Figure 1. The error bars were computed from the conventional statistical equations with multiple measurements at selected points. For small CO_2 concentrations of entering air streams and assuming constant total gas flow rate across the entire column, the percent CO_2 absorbed from air was computed for both figures as following:

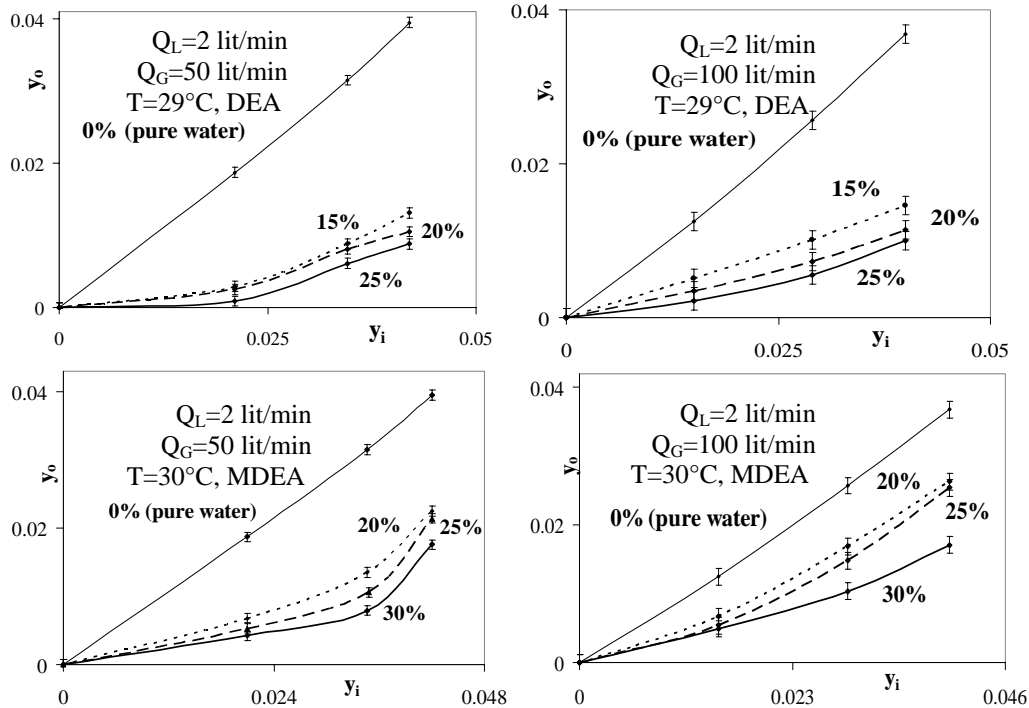$$\% \, (CO_2) \, Absorbed = \frac{y_{inlet} - y_{outlet}}{y_{inlet}} \times 100 \qquad (1)$$

Figure 1 – Measured output gas concentrations versus mole fractions for DEA and MDEA solvents at corresponding operating conditions.
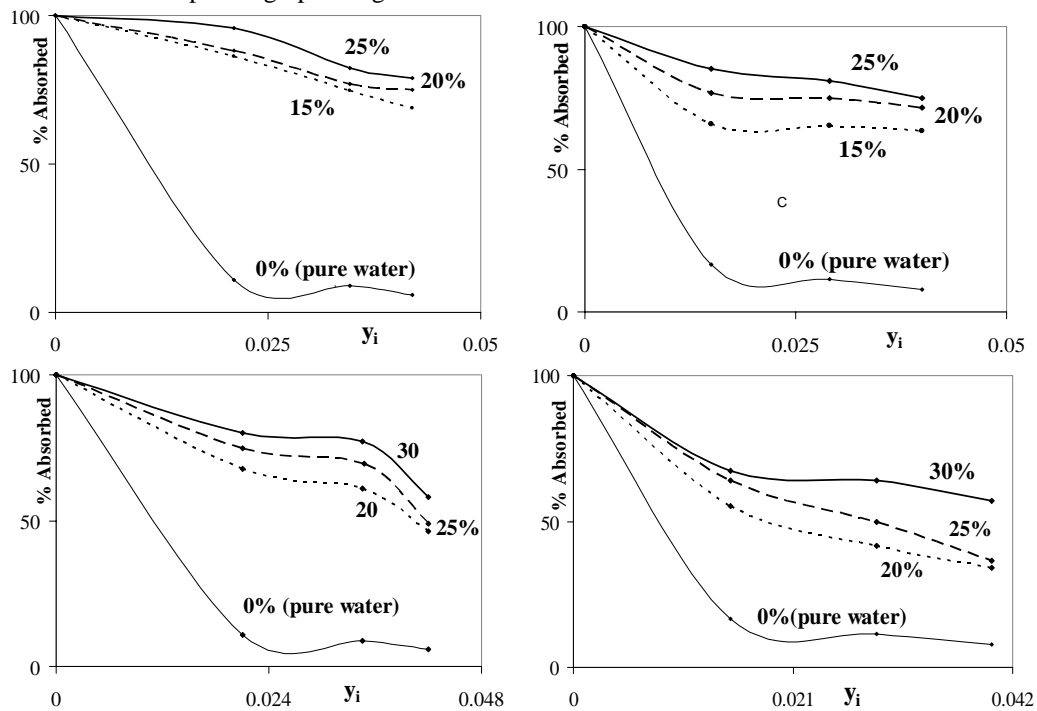

Figure 2 – Calculated values for percents of absorbed CO2 versus input concentrations for various solvents and operating conditions mentioned in Fig. 1 respectively.

Evidently, the outlet gas concentration increases by increasing the inlet CO_2 mole fraction at constant solvent concentration as shown in Figure 2. On the other hand, increasing the CO_2 mole fraction of entering air (provided that all other conditions are kept constant) although increases the absorption rate (due to relatively larger driving forces), however it may not necessarily lead to higher percents of CO_2 absorption. Actually, in most cases, the increase in mass transfer rate is much smaller than the increase in inlet concentration due to limitations of equilibrium concentrations[1] (or limited solvent capacity). In such cases, the percent of CO_2 absorption will decrease with increase in the CO_2 mole fraction of entering air.

Evidently, the absorption capacity of the solvents increases by increasing the DEA or MDEA concentrations of the entering solutions. Furthermore, since DEA is much stronger alkali compared to MDEA, hence, DEA solutions are able to absorb more carbon dioxide than MDEA solutions (at constant solvent concentrations). The collected experimental data were used to simulate the absorption process by employing various artificial neural networks.

3. A brief overview of neural networks

Neural networks generally consist of several interconnected neurons in one or more *hidden* layers. They can be classified from different points of views such as the type of input transformation, their structural architecture and the type of learning algorithm. Neural networks may employ either projection or kernel based transformations to account for correlation among the inputs.

In the first transformation the inputs are projected on a single axis, the projection may be linear or non-linear. The McCulloch-Pitt neuron, Perceptron and Adaline are examples of linear projections [12]. In the second transformation, the norm (usually Euclidean) of the input vector with respect to a fixed point (centre) is used. Radial basis function networks are the most popular examples of the kernel based input transformations [13].

In its simplest form, a feed-forward network is constructed from an *input layer* of source nodes that are projected onto an *output layer* of computation nodes via synaptic weights. The "single layer" designation refers to the output layer containing feed-forward computation nodes (neurons). A linear associative memory is an example of a single-layer neural network. The network associates an output pattern (vector) with an input pattern (vector), and the information is stored in the network by virtue of the modifications made to the synaptic weights of the network.

Multi-layer feed-forward neural networks contain one or more *hidden layer(s),* whose corresponding computational nodes are called *hidden neurons* or *hidden units*. The function of the hidden neurons is to intervene between the external input and the network response. By adding one or more hidden layers, the network is enabled to extract higher order statistics (more information) by virtue of the extra set of synaptic connections and increased neural interactions. The ability of the hidden neurons to extract higher order statistics is crucial for large input dimensions. The neurons of each layer may be either *partially* or *fully* connected to the neighboring layers.

The learning algorithm of a neural network deals with the adjustment of the network parameters and usually settles to solving an unconstrained or constrained optimization problem. In many engineering applications we are concerned with the estimation of an underlying trend (or function) from a limited number of input-output data points with little or no knowledge of the form of the true function (truth). The set of examples (the *training set*) contains elements that consist of paired values of the independent (inputs) and the dependent (outputs) variables. A supervised learning algorithm adjusts the network parameters according to the differences between the measured response $y(\underline{x}_i)$ and the network outputs $\hat{y}(\underline{x}_i)$ corresponding to a given input $\underline{x}_i$. Supervised learning requires a supervisor, to provide the target signals.

3.1. Training of the MLP network

The basic element of a Multi Layer Perceptron (MLP) neural network is the artificial neuron which performs a simple mathematical operation on its inputs. The input of the neuron consists of the variables $x_1,.....,x_p$ and a threshold (or bias) term. Each of the input values is multiplied by a weight, w_i, after which the results are added with the bias term to produce z. Finally, as shown in Figure 3, a known activation function, φ, performs a pre-specified (non-linear) mathematical operation on the projected inputs. Various activation functions such as sigmoid or hyperbolic tangent are traditionally used for this purpose [13].

---

[1] Note that the equilibrium constant is actually a function of solute concentration in both phases as well as the operating temperature and pressure.

The MLP network is trained by adapting the synaptic weights using a back-propagation technique or any other optimization procedure. During training phase, the network output is compared with a desired output. The error between these two signals is used to adapt the weights. This rate of adaptation may be controlled by a learning rate. Additional linear weights ($\alpha's$, as shown in Fig. 3) were used in this work to accelerate the network convergence. The optimal values of these linear parameters were updated after each back-propagation iteration using the following set of linear equations:

$$\left(\Phi^T \Phi\right)\underline{\alpha} = \Phi^T \underline{y} \tag{1}$$

where $\Phi_{i,j} = \varphi(z_{i,j})$, $i = 1,...,N$ & $j = 1,....,M$ and y is the $N \times 1$ vector of measured values. The parameters N and M represent number of training data and number of neurons respectively. The training flow chart of such MLP network is given in our previous articles [13].

3.2. Training of the RBF network

The training of a projection based network (such as MLP) always reduces to the solution of a large-scale non-linear optimization problem. Such problems are usually very time demanding and often encounter severe convergence problems. In contrast, the training of RBF networks with pre-specified non-linearities (centers and spreads) reduces to the solution of an over-determined set of linear equations which can be solved by a variety of highly stable techniques. An efficient in-house algorithm is used to effectively train the so called Regularization or Radial Basis Function (RBF) neural networks with isotropic spreads.
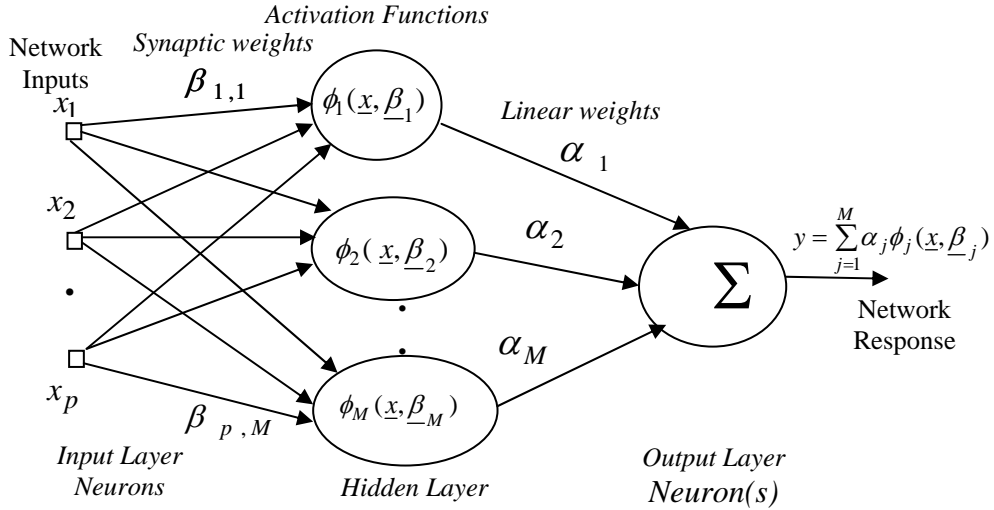

Figure 3 – Schematic representation of MLP and RBF feed-forward neural networks.

4. Simulation results

The entire collection of experimental data for CO_2 outlet concentrations and percentage of absorbed CO_2 were used to train both MLP and RBF networks. By definition, the RBF regularization network employs the same number of neurons as the data points. For better comparison of both networks performances, the number neurons of MLP networks were kept equal to the number of training exemplars.

The regularization network was completely self-sufficient and did not require any initial values for its linear and non linear parameters. As it was mentioned in the theoretical section, these networks use all the input data as their centers. Furthermore, our in-house optimization technique [12-15] was used to select the optimal value of isotropic spreads at each case, separately. The leave one out cross validation (LOOCV or CV in brief) criterion was also employed to provide the optimum level of regularization for a given set of centers and the optimal isotropic spread.

In contrast, the MLP networks performances completely depended on the set of initial values selected for their synaptic weights. In this work, various initial synaptic values were tried for each training data set and then the best values were selected based on visual considerations. Evidently, the above procedure for training of MLP networks is much more time demanding than the previously described [15] straightforward method for training of Regularization networks.

Figures 4 and 5 compare typical performances of optimally regularized RBF networks with the best selected MLP network for prediction of CO_2 outlet concentrations and percents of CO_2 absorption versus CO_2 mole fractions of the entering air streams for the experimental data depicted in Figures 1 and 2. Although both networks performs very close to each other, however, the RBF network provides

slightly better performances with less oscillations due to its noise filtering capability. Furthermore, as it was mentioned above, the appropriate choice of initial synaptic weights was crucial for the proper performance of MLP network, while RBF does not need any initial values for its synaptic weights.
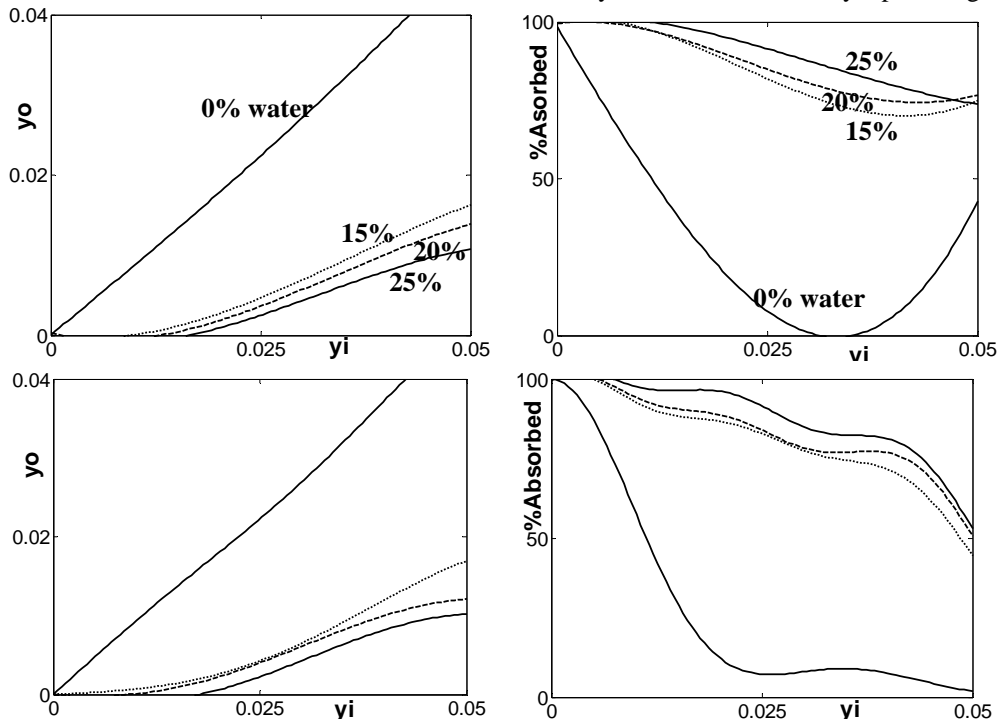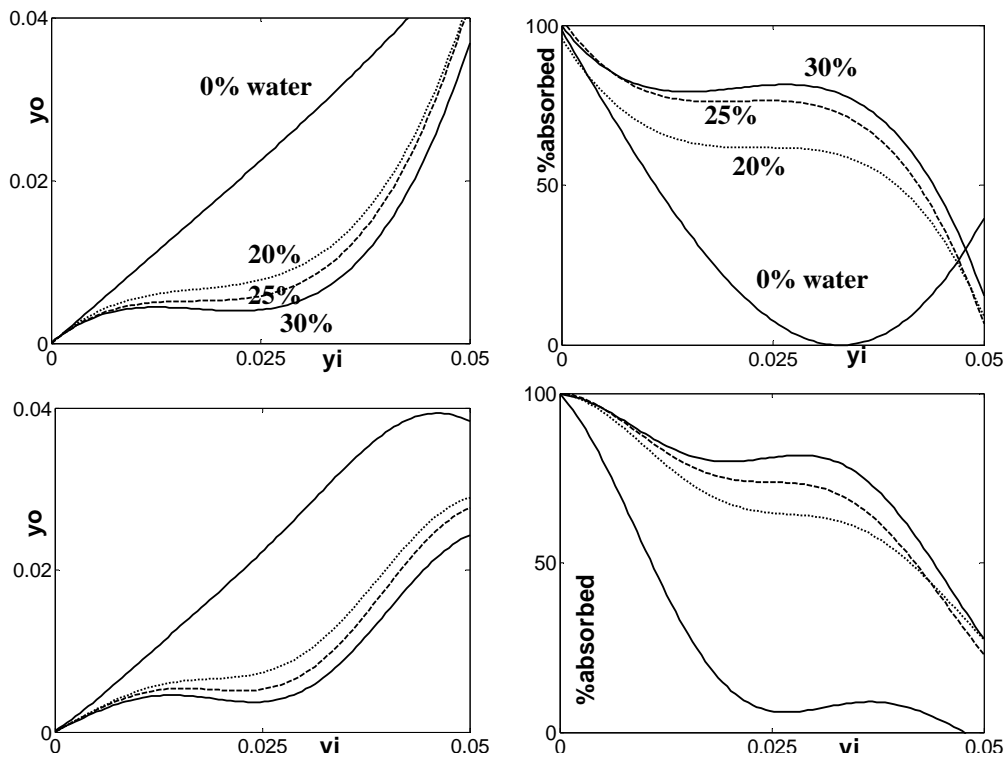

Figure 4 - Typical comparison and predicted values by MLP network (top) with RBF network (bottom) for outlet carbon dioxide concentration and percentage absorption. (Q_G=50lit/min, Q_L=2 lit/min, <u>DEA solvent</u>)

Figure 5 - Typical comparison of predicted values by MLP network (top) with RBF network (bottom) for outlet carbon dioxide concentration and percentage absorption. (Q_G=50lit/min, Q_L=2 lit/min, <u>MDEA solvent</u>).

5. Conclusion

The experimental data collected in this article can be used as an influential tool to improve our understandings of the packed absorption processes especially the absorption of CO_2 from air by various alkanolamine solutions. The simulation results presented here show that both MLP and RBF networks can perform adequately for prediction of outlet CO_2 concentrations or the percent of absorption. The RBF network with isotropic spreads does not need initialization and usually perform more appropriately due its solid theoretical background (multivariate linear regularization theory) and its noise filtering capability.

6. Acknowledgement

7. References

1. XU X., SONG Ch., Miller B. G.; Scaroni A.W., Xu. Adsorption separation of carbon dioxide from flue gas of natural gas-fired boiler by a novel nanoporous 'molecular basket' adsorbent. Fuel processing technology. 86 (2005) 14-15,285.
2. Fauth, D.J, Frommell, E.A, Hoffman, J.S, Reasbeck, R.P and pennline, H.W. Eutectic salt promoted lithium zirconate: Novel high temperature sorbent for CO_2 capture. fuel processing technology, 86(2005)1503_1521
3. Gray, M.L., Soong, Y., Champagne, K.J., Pennline, H., Baltrus, J.P., Stevens Jr, R.W., khatri, R., Chuang, S.S.C., Filburn, T. Improved immobilized carbon dioxide capture sorbents. fuel processing thecnology,86(2005)1449-1455
4. Park, S.W., Suh, D.S., Hwang ,K.S. and Kumazawa, H.Gas absorption of carbon dioxide in a hollow fiber cotained liquid membrane absorber.chemical engineering, 14(4) (1997) 285-291
5. Datta, A.K.,Sen,P.K.,Optimization of membrane unit for removing carbon dioxide from natural gas.journal of membrane science, 283 (2006) 291-300
6. Yeh, J.T., Pennline, H.W., Resnik, K.P., Study of CO2 absorption and desoorption in a packed column. Energy and Fuels, 15 (2001)274-278.
7. Vaidya, P. D., Kenig, E. Y., Absorption of CO_2 in to aqueous blends of alkanolamines prepared from renewable resources. Chemical Engineering Science, 62 (2007)7344-7350
8. Huttenhuis, P. J. G., Agrawal, N. J., Hogendoorn, J. A.,Versteeg, G. F. Gas solubility of H_2S and CO_2 in aqueous solution of N-methyldiethanolamine. Journal of Petroleum Science and Engineering, 55 (2007)122-134.
9. Lin Sh. H., Shyu. Ch. T., Performance characteristics and modeling of carbon dioxide absorption by amines in a packed column. Waste Management 19 (1999)255 – 262.
10. Sultan, G.I., Hamed, Ahmed M. *, Sultan, A.A., The effect of inlet parameters on the performance of packed tower-regenerator. Renewable Energy, 26 (2002)271–283
11. Sharma, R., Singh, K., Singhal, D., Ghoshd, R., Neural network applications for detecting process faults in packed towers. Chemical Engineering and Processing , 43 (2004) 841–847
12. Shahsavand, A., Optimal and adaptive radial basis function neural networks, Ph.D. Thesis, University of Surrey, UK, 2000.
13. Shahsavand, A., Pourafshari Chenar, M. Neural networks modeling of hollow fiber membrane processes. Journal of Membrane Science, 297 (2007) 59-73
14. Shahsavand, A., Ahmadpour, A., Application of optimal RBF neural networks for optimization and characterization of porous materials. Computers and Chemical Engineering, 29 (2005)2134-2143
15. Shahsavand, A., An Optimal Radial Basis Function (RBF) Neural Network for Hyper-surface Reconstruction.Scintia iranica, accepted for publication. (2009)