International Conference on Chemical & Bioprocess Engineering, Kota Kinabalu, August 27-29 03, pp 755-760

PORE SIZE DETERMINATION OF ASYMMETRIC MEMBRANE USING NEURAL NETWORK

Khairiyah Mohd. Yusof[a] **Ani Idris**[b] **Lim, J. S.**[c]

[a]*Chemical Engineering Department, Faculty of Chemical and Natural Resources Engineering, 81310 UTM, Skudai, Johor Bahru, Malaysia*
Tel: 607-5535612, Email: k-khairi@fkkksa.utm.my

[b] *Bioprocess Engineering Department, Faculty of Chemical and Natural Resources Engineering, 81310 UTM, Skudai, Johor Bahru, Malaysia*
Tel: 607-5535604, Email: ani@fkkksa.utm.my

[a]*Chemical Engineering Department, Faculty of Chemical and Natural Resources Engineering, 81310 UTM, Skudai, Johor Bahru, Malaysia.*

Abstract
This study, investigates the possibility of applying artificial neural network (ANN) as an alternative method to estimate the pore size of the asymmetric hollow fiber membranes. ANN, a connectionist-based (black box) model, consists of layers of nodes with nonlinear basis functions and weighted connections that link the nodes. Using the nodes and weights, the inputs are mapped to the outputs after being trained with a set of training data. The input data needed for training the ANN model, the solute rejection and the permeation rate, are obtained from permeation experiments. Since the number of experimental data points needed for training the ANN model is limited, stacked neural network is utilized instead of the more common and simple feedforward ANN. With the development of this ANN model, the procedure to estimate membrane pore size was found to be easier and faster with a testing error of less than 2% compared to the experimental data.

Keywords
Asymmetric membranes, stacked network, artificial neural network, pore sizes

Introduction
Neural networks have been attracting great interest as predictive models, as well as for pattern recognition (Nascimento, Giudici and Guardani, 2000). For the past two decades, neural networks are a rapidly growing field of artificial intelligence that has found many applications in process modeling especially for nonlinear systems (Hao, Youngseok and En, 1998).

Artificial neural networks (ANN) have the ability to map non-linear relationships without prior information about the process or system model. Their advantages over the classical mathematical models are the simultaneous identification of structure parameters as well as the ability to "learn" and adapt by examples. ANN modelling, however, is no replacement for a good understanding of process behaviour but it makes possible to develop quickly models of complex reactions (Morris, Montague and Willis, 1994). The success in obtaining reliable and robust network depends strongly on the choice of process variables involved, as well as the available set of data and the domain used for training purposes (Nascimento, Giudici and Guardani, 2000).

Various methods have been used to characterize the pore size and pore size distribution of hollow fiber membranes. The 3 general methods generally used are i) microscopy observation method, ii) thermoporometry method and iii) permeation experiments. In cases of asymmetric hollow fiber membranes, due to the very small pore sizes, microfiltration characterization techniques such as microscopy observation method cannot be used for asymmetric membranes (Mulder , 1981). Generally, the last method is used, based on permeation and rejection performance using reference molecules and particles. In order to characterize the pore size and pore size distribution from permeation experiments, quantitative transport model and the so called "pore model" or hydrodynamic model has been used. In recent years the surface force pore flow model has been developed to correctly characterize the pore size and pore size distribution (Ani *et al,* 2002 and Tam *et. al*.1993). It contained the interfacial force parameters and friction force parameters and the effect of concentration polarization and operating conditions. However, the mathematical solutions involved in the determining the pore size and pore size distribution involved complicated equations, with trial and error calculations that is complicated and tedious. Thus, in this study, artificial neural network is applied as an alternative method to estimate the pore size of the asymmetric hollow fiber membranes. The raw data needed for the artificial neural network are obtained from permeation experiments, that is the solute rejection and the permeation rate. With the development of this neural network model, the procedure to estimate membrane pore size could be made easier and faster with an output error of less than 2% compared to experimental data.

In reviewing literature on membrane characterization, so far there has not been any work found that uses ANN for modeling. A hindering factor is the small number of data points available for training from permeation experiments. In this research, this obstacle is overcome by utilizing stacked network instead of the more commonly used feed forward ANN.

Conventional Methods to Estimate Pore Size

Asymmetric membrane can be considered as porous where the thin top layer of the membrane supported by a porous sub layer, with the resistance to mass transfer being almost completely determined by the top layer. For this reason, the characterization of asymmetric membranes involves the characterization of the top layer such as its thickness, pore size distribution and surface porosity. Because of the small pore sizes, microfiltration characterization techniques cannot be used for asymmetric membranes. The resolution of an ordinary scanning electron microscope is generally too low to determine the pore sizes in the top layer accurately. Furthermore, mercury intrusion and bubble-point methods cannot be used because the pore sizes are too small, so that very high pressures would be needed, which will destroy the polymeric structure.

Two general methods to estimate pore sizes of asymmetric membrane are briefly discussed here (Mulder, 1996):

a) Thermoporometry
b) Permeation experiments

Thermoporometry

Thermoporometry is based on the calorimetric measurements of a solid-liquid transition (e.g. of pure water) in a porous material and can be applied to determine the pore size in porous membranes. This may be the pores in the skin of an asymmetric membranes, the temperature at which the water in the pores freezes (the extent of undercooling) depending on the pore size. As the pore size decreases the freezing point of water decreases. Each pore (pore size) has its own specific freezing point.

For cylindrical pores containing water, the following equation for melting can be derived:

$$r_p = 0.68 - \frac{32.33}{\Delta T} \qquad (1)$$

where r_p is the pore radius (nm) and ΔT the extent of undercooling (˚C). It can been seen from equation (i) that as the pore radius becomes smaller the extent of undercooling increases.

Permeation Experiments

This method is based on permeation and rejection performance using reference molecules and particles. In order to characterize the pore size and pore size distribution for permeation experiments, quantitative transport model such as the surface force pore flow (SF-PF) model has been used by many researchers (Ani et al, 2002, Tam et. al, 1993, Matsuura and Sourirajan, 1981). The SF-PF model was first suggested by Matsuura and Sourirajan (1981) and since then it has been used by many researchers to describe and predict the performance of reverse osmosis type membranes. Matsuura and Sourirajan (1981) developed general expressions based on the preferential sorption-capillary flow mechanism, taking into account the surface forces acting on the solute to analyse experimental reverse osmosis data. The two main equations used for this model, the solute separation equation and momentum balance equation are shown below:

$$f' = 1 - \frac{\int_0^1 \left[\exp[\alpha(\rho)]/1 + \left[b(\rho)/e^{-\Phi(\rho)}\right](\exp[\alpha(\rho)]-1)\right](\alpha(\rho)\rho d\rho}{\int_0^1 \alpha(\rho)\rho d\rho} \qquad (2)$$

$$0 = \frac{\partial^2\alpha(\rho)}{\partial\rho^2} + \frac{1}{\rho}\frac{\partial\alpha(\rho)}{\partial\rho} + \frac{\beta_2}{\beta_1} + \frac{1}{\beta_1}(1 - e^{-\xi(\rho)})[C_A(\rho) - 1] - \frac{[b(\rho)-1]\alpha(\rho)C_A(\rho)}{\beta_1} \qquad (3)$$

These two complicated equation have to be solved by trial and error, which is rather complicated and lengthy. Although this problem can be solved using conventional simulation method it involved many iterations before the final solution obtained. Thus in this study, artificial neural network is applied as an alternative method to estimate the pore size of the asymmetric hollow fibre membranes.

Artificial Neural Network Model Architecture

Artificial neural networks (ANN) are made up of individual models of biological neuron that are connected together to form a network. Unlike the human brain which consists of between 10^{11} to 10^{14} neurons, the number of artificial neurons in an ANN is relatively very small. The connections within ANN are much more simplified than that of the brain (Muller and Reinhardt, 1990). However, despite the simplicity, an ANN designed using only a few neurons can be trained to learn information pattern recognition and systems indentification.

An ANN is constructed of interconnected basic elements called neurons or nodes. A typical feedforward neural network model constitutes three types of layers of nodes. The first layer is known as input layer. This layer receives information from an external source and passes the information to the second layer of nodes, which is known as the hidden layer. There can be more than one hidden layer. However, one hidden layer is normally sufficient for chemical and biochemical processes. The last layer of nodes, the output layer, receives processed information from the network and sends the results an external receptor (Baughman and Liu, 1995). Figure 1 illustrates a typical three-layer network with 4 nodes in the hidden layer.

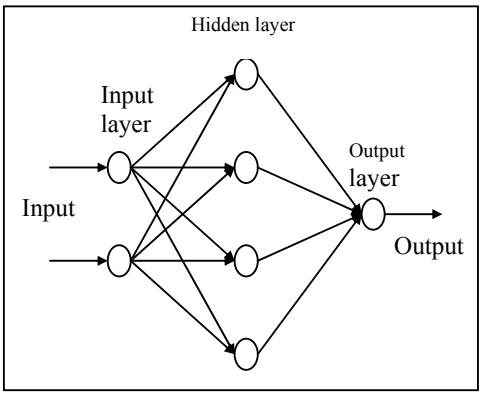

Figure 1 – A feedforward network with 4 nodes in the hidden layer.

At present, the most commonly used ANN is the feedforward backpropagation network, which accounts for most neural network applications (Hao, et al., 1998). Multilayer feedforward ANN were mathematically proven to be a universal approximator [Hornik, et. al., 1989].

To develop a neural network model, three phases are required; training phase, recall phase and generalization phase. In the training phase, a set of training data is repeatedly presented to the neural network model to teach it. The weights between the nodes are optimised until the specified input yields the desired output. Through these adjustments, the neural network "learns" the correct input-output behavior. In the recall phase, input patterns from the training data will be used to test the networks and adjustments will be made to make the system more reliable and robust. During the generalization phase, input patterns that have not been seen in training will subjected to networks model in order to monitor the system performance (Baughman and Liu, 1995).

Stacked Neural Network

Using a single basic neural network with few experiment data points, it is difficult to guarantee that a good predictive model will be obtained in the complete experimental domain (Lanouette et al, 1999). To improve the accuracy of a model when only a limited number of experimental data points in the training data set is available, stacked neural network is recommended (Wolpert, 1992). In stacked neural networks, several different neural network models are combined in order to improve model performance. Since each neural network representation can behave differently in different regions of the input space, representational accuracy over the entire input-output space can be improved by combining several neural network models (Zhang et al., 1997). A sample architecture for a stacked neural network is shown in Figure 2.

Starting from the identical training data set, a large number of different neural network models can be obtained, using each time a different set of initial weights or using a different subset of the training data set (Lanoutte et al., 1999). The outputs of these networks, called the level-0 models outputs, along with the original input data, are then used as inputs to other models, at a higher level in the stacking structure. The second level of model, called the level-1 model, is developed using the results of level-0 model (Ali *et al*, 2001).

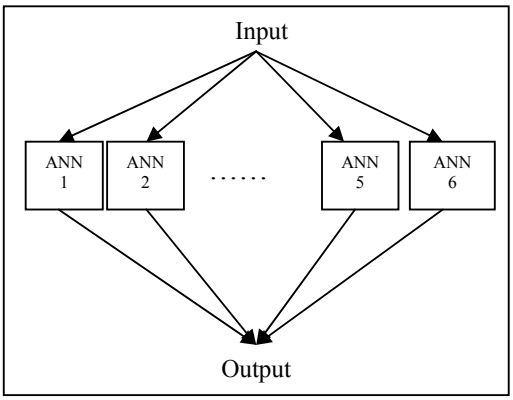

Figure 2 – Architecture of a stacked network.

A set of training data set is needed in order to develop the stacked neural network. Firstly, let us denote level-0 and level-1 data set as D_{L0} and D_{L1} respectively. For level-0, a few networks models have to be developed using data set D_{L0}. Next, one experimental data point is removed from D_{L0}. This subset is known as D_1. By using the same method, more different subset of data D_i can be obtained by removing different data points. Each of these subsets is then used as training data for the candidate ANN models that has been developed earlier. The output predicted by these models is combined with the actual output to develop level-1 models.

There exist numerous types of level-1 model to produce the output stacked model. A simple approach is to take equal weights for the individual networks (Zhang *et al*, 1997). A second way is to obtain a weighted sum of each prediction, the weight corresponding to inverse of the contribution of each level to the sum of squares of the errors calculated which is known as weighted output (Lanouette *et al*, 1999). The third way is to combine the models is by using principal component analysis (PCA). Wolpert (1992) an alternative method, where the outputs from level-0 models are used as training data to train a new level-1 ANN model. The output from this model is the final output for the stacked network.

Model Development

In this work, all ANN models were developed in MATLAB environment and utilized MATLAB neural network toolbox. All ANN models had one hidden layer with the sigmoid function as the activation function. The models were trained using backpropagation algorithm with the

Levenberg-Marquadt method used to optimize the weights and biases. To develop the ANN models, the number of nodes in the hidden layer and the maximum acceptable error during training were varied. The number of training and testing data was also varied. The results given in this paper are the best found during the study.

Evaluations of the models are based on root mean squared (RMS) error from each model prediction of the test data. Error is defined as the difference between desired (or actual value provided by the testing data) output value and the predicted output value.

Ten individual neural network models had been developed to estimate pore size with different subset of training data and network architecture. There were 6 raw experimental data, as shown in Table 1. Data number 5 was used for the overall stacked model validation, and the rest were used for the individual ANN model development in level-0. To create the training subset for level-0 models, 4 data points were assigned as the training data, while 1 point was used as testing data to validate the individual model. For example, ANN1, which has 4 nodes in the hidden layer, was trained using data number 2, 3, 4 and 6 until the weights and biases have been optimised to give an acceptable training error. It was then validated with the test data in the generalization phase for ANN1 using data number 1. ANN2, which has 6 nodes in the hidden layer, had the same training and testing data as ANN1. Each odd-numbered network has 4 nodes and each even numbered network has 6 nodes in the hidden layer. Each pair of network from ANN3 to ANN10 was similarly formulated, except different training and testing subsets were utilised.

For the level-1 model, three main techniques have been used in this study to combine the individual level-0 neural network models. They are output averaging, weighted average, and a level-1 ANN model.

Table 1: Experimental data for the system NaCl-H_2O using porous cellulose acetate membrane (Matsuura and Sourirajan, 1981).

	Permeate Rate, PR x 10^{10}kg/hr	Solute Separation, f	Membrane pore size, R x 10^{10}m
1	69.45	0.978	6.9
2	69.63	0.965	8.2
3	90.19	0.885	12.7
4	112.13	0.857	13.2
5	133.67	0.827	13.7
6	156.94	0.679	18.2

Results

Table 2 gives the calculated root mean square error (RMSE) for single network predictions based on different training and testing data. The first column shows the training data number that was used for training. For example, when data numbers 2, 4 and 6 were used for training, data numbers 1, 3 and 5 were used for testing (the RMSE for this testing set is 0.3208).

The RMSE of the individual ANN models in level-0 of the stacked network are given in Table 3. RMSE for stacked network with different methods of combination are shown in Table 4.

Table 2 - Single network prediction.

Training data set	Testing data set	RMSE
2,4,6	1,3,5	0.3208
1,2,4,6	3,5	0.2563
1,2,3,4,6	5	0.1469

Table 3 - RMSE of level-0 ANN models.

ANN	RMSE
ANN 1	0.57576
ANN 2	0.70199
ANN 3	0.34749
ANN 4	0.01156
ANN 5	0.90903
ANN 6	0.27073
ANN 7	0.26749
ANN 8	0.81604
ANN 9	0.27714
ANN 10	0.61566

Table 4 - Stacked network prediction with different level-1 models.

Method	RMSE
ANN	0.0078
Average	0.3877
Weighted average	0.1384

Discussion

From Table 2, the results show that the best single network prediction, 0.1469, was obtained when 5 data points were used for training and 1 for testing. For the stacked network model results shown in Table 4, although the level-0 model predictions were unsatisfactory, the generalized stacked network (ANN in level-1) provided the best performance with an error of less than 1%. This is a big improvement in predicting the test data compared to the individual single network models.

Among the three level-1 models, using a straight-forward average gave the worst prediction. The prediction RMSE is also greater than the single network predictions. However, the weighted average level-1 model prediction is better than the single network prediction. This is expected,

because taking a simple average does not give priority to the level-0 model with the better prediction. The weighted average performed better because the connection weights from the corresponding network output was taken into account. Using another ANN in level-1 gave the best prediction because the network was able to learn and improve the prediction further during the learning phase.

Conclusion

In this study, limited experimental data are insufficient for development of ordinary feedforward network model. Ten networks have been developed in order to provide more data to increase the accuracy of stacked network. The combination of the ten models has shown a better result compared to the performance of individual neural networks. A general conclusion form this study is that the stacked networks developed is able to predict the size of pore size with low percentage of error. Although the network can only be verified with one testing datum out of the six raw experimental data available, the results show that ANN is able to model the pore size prediction. Based on the results obtained, the networks are reliable and could be an alternative to estimate pore sizes in the future.

For future work, estimation of pore size can be carried out for other types of membrane besides asymmetric membrane. In addition, a higher number of data points is desirable to obtain a better model and verification of the model.

References

Ani Idris, Ismail A. F.,. Shilton S. J, Roslina, R., Musa, M (2002). The Deduction of Fine Structural Details of Reverse Osmosis Hollow Fiber Membranes Using Pore Modelling. *Separation and Purification Tech.* (29) 217-227.

Ali A. Ghorbani and Kiarash Owrangh (2001). Stacked Generalization in Neural Networks: Generalization on Statistically Neural Problems. *International Joint Conference on Neural Networks (IJCNN'2001)*, 1715-1720

Baughman, D.R. and Liu, Y.A, (1995) Neural Networks in Bioprocessing and Chemical Engineering. *Academic Press*, 1-58.

Hao, W., Youngseok, O. and En, S.Y. (1998). Strategies for Modeling and Control of Nonlinear Chemical Processes Using Neural Networks. *Comp. Chem. Eng. Vol*. 22, 823-827.

Hornik, K., Stinchcombe, M., and White, H. (1989). "Multilayer Feedforward Neural Networks are Universal Approximators", Neural Networks, vol. 2, 359-366.

Lanouette, R., Thibault, J. and Valade, J.L. (1999). Process Modeling with Neural Networks Using Small Experimental Datasets. *Computers and Chemical Engineering*. Vol.23, 1167-1176.

Matsuura, T. and Sourirajan, S. (1981). Reverse Osmosis Transport through Capillary Pores Under the Influence of Surface Forces, Ind. Eng. Chem. Processes Des. Dev., 20 (1981) 273 – 282.

Morris, A.J., Montague, G.A. and Willis, M.J. (1994) Artificial Neural Networks: Studies in Process Modelling and Control. *Trans.I. Chem. Eng. 72A*, 3-19

Mulder, M. (1996). Basic Principles of Membrane Technology. *Kluwerr Academic Publishers*, 157-189.

Muller, B. and Reinhardt, J. (1990). Neural Networks an Introduction. *Springer-Verlag*, 2-22.

Morris, A.J., Montague, G.A. and Willis, M.J. (1994). Artificial Neural Networks: Studies in Process Modeling and Control. Trans. I. Chem. E. 72A, 3-19.

Nascimento, C.A.O., Giudici, R. and Guardani, R. (2000). Neural Network Based Approach for Optimization of Industrial Processes. Comp. Chem. Eng. Vol. 24, 2303-2314.

Sridhar, D.V., Seagrave, R.C. and Bartlett, E.B. (1996). Process Modeling Using Stacked Neural Networks. *AIChE Journal* Vol.42, 2529 – 2539.

Tam, C. M., Matsuura, T. and Tweddle, T.A. (1993). Polysulfone Membrane: iii Performance Evaluation of PES-PVP Membrane, Sep. Science and Tech., 28 2621 – 2633.

Wolpert, D.H. (1992). Stacked Generalization. *Neural Networks* Vol. 5, 241-259.

Zhang, J., Martin, E.B., Morris, A.J. and Kiparissides, C. (1997). Inferential Estimation of Polymer Quality Using Stacked Neural Networks. *Computers and Chemical Engineering*. Vol.21, 1025-1030.

Nomenclature

C_A	dimensionless solute concentration at the pore outlet
f'	true value of solute separation by membrane pore
r	radial distance, m

Greek Letters

ρ	dimensionless quantity defined by eqn. (a)
$\alpha(\rho)$	dimensionless solution velocity profile in the pore
β_1	dimensionless solution viscosity
β_2	dimensionless operating pressure:
α	activity coefficient
$\Phi(\rho)$	dimensionless potential function:
ϕ	potential in the interfacial force field
$b(\rho)$	dimensionless friction function

Subscripts

1	bulk feed solution on the high pressure side of membrane
2	concentrated boundary solution on high pressure side of membrane

Subscripts

1	bulk feed solution on the high pressure side of membrane
2	concentrated boundary solution on high pressure side of membrane
3	membrane permeated product solution on the low-pressure side of membrane