Feed Forward Neural Network Model for Isopropyl Myristate Production in Industrial-scale Semi-batch Reactive Distillation Columns

Nur Alwani Ali Bashah,[1,2] Mohd Roslee Othman[2] and Norashid Aziz[2*]

[1]Faculty of Chemical Engineering, Universiti Technologi Mara,
40450 Shah Alam, Selangor, Malaysia
[2]School of Chemical Engineering, Universiti Sains Malaysia, Engineering Campus,
14300 Nibong Tebal, Pulau Pinang, Malaysia

[*]Corresponding author: chnaziz@eng.usm.my

Abstract: *The application of the artificial neural network (ANN) model in chemical industries has grown due to its ability to solve complex model and online application problems. Typically, the ANN model is good at predicting data within the training range but is limited when predicting extrapolated data. Thus, in this paper, selected optimum multiple-input multiple-output (MIMO) and multiple-input single-output (MISO) models are used to predict the bottom (x_b) compositions of extrapolated data. The MIMO and MISO models both managed to predict the extrapolated data with MSE values of 0.0078 and 0.0063 and with R^2 values of 0.9986 and 0.9975, respectively.*

Keywords: Feed forward neural network, extrapolation, semi-batch reactive distillation, industrial scale, normalisation

## 1.	INTRODUCTION

Batch reactive distillation (BRD) is a dynamic process that operates under unsteady conditions. By coupling reaction and distillation in a single unit, the model results in a large number of complex differential equations. Thus, the development of the model is expensive and time consuming to solve. Due to that, it may not be practical to use the complex differential equations to develop the BRD process. One alternative method for representing this process is to develop an empirical model. Artificial Neural Networks (ANNs) are one type of available empirical model that can be used to solve various types of mathematical problems in BRD, such as modelling,[1,2] control systems,[3] soft sensors,[4,5] a combination of soft sensors and control,[6,7] and optimisation.[8]

Recently, the BRD model has gained attention for use in esterification and transesterification processes. Extensive literature is available regarding fundamental models of the esterification process, such as ethyl acetate, butyl acetate, methyl acetate, and the hydrolysis of lactic acid and isopropyl acetate.

However, the transesterification modelling process, especially for long chain fatty esters in batch reactive distillation, is limited. Li et al.[8] developed a thorough fundamental model for isopropyl myristate (IPM) production in BRD. However, this model requires a large number of equations. Thus, shifting to the ANN model will simplify and expedite model convergence. In addition, the ANN model is advantageous for real time applications, such as control systems.

The feedforward neural network (FFNN) consists of three layers, namely input, hidden and output layers. Prior to developing the FFNN model, the input-output data with variables of different magnitudes must be scaled. The z-score normalisation technique is yet to be tested in the BRD process. On the other hand, the min-max normalisation technique is commonly used for BRD processes, as demonstrated by Bahar et al.[4] and Konakom et al.,[6] and is significantly effective for modelling the training range. In this work, a multiple-input multiple-output model (MIMO) and a multiple-input single-output model (MISO) were developed to predict bottom composition (x_b). The developed MIMO and MISO models were both used to test their abilities for predicting extrapolated data and to test their performance in terms of their mean square errors (MSE) and R^2 values.

2.   EXPERIMENTAL

2.1   Process Description

The transesterification model of methyl myristate and isopropanol in industrial scaled semi-batch reactive distillation (BRD) is simulated based on the work conducted by Seader et al.[9] The simulation was conducted using the commercial simulator Batchfrac. All process variables were designed in Batchfrac to achieve 98% distillate purity and the complete conversion of methyl myristate. The distillation column consisted of 30 trays, with a total condenser and a reboiler. Prior to the simulation, the column pressure was computed using the procedure adopted by Jimoh et al.[10] It was assumed that the reversible reaction shown in Equation 1 only occurred in the reboiler, and the reaction kinetics were obtained from Bashah et al.[11]

$$C_{15}H_{30}O_2 + C_3H_8O \leftrightarrow C_{17}H_{34}O_2 + CH_4O \tag{1}$$
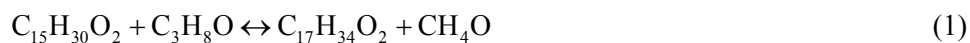
The initial batch consisted of pure methyl myristate (MM), isopropanol (IP), methanol (M) and a homogeneous catalyst. The products formed from this reaction included isopropyl myristate (IPM) and methanol (the most volatile component). These products were collected in a reboiler and accumulator,

respectively. After a specified reaction period, the limited flow of IP was continuously fed into the reboiler. Next, the results from the simulation were compared with the results obtained from Seader et al.[9] When comparable results were achieved, the Batchfrac model in Aspen Plus® was used for the sensitivity study and for generating data for developing the ANN model.

2.2   Neural Network Model Development

The ANN model learns from the relationships between the input and output data. A few sets of input-output data for developing the ANN model were generated by the validated Batchfrac model.

2.2.1   Pre-processing and division of data

All of the raw data were of different magnitudes. Thus, the larger magnitude variables would be dominant over the smaller variables. Thus, the data must be normalised before model development. For this purpose, the z-score normalisation technique is used based on the mean and standard deviation of the given data, as shown in Equation 2. This method is useful when the minimum and maximum values of the variables are unknown. After the data are normalised, all of the set data are divided into training, validation testing and extrapolation data.

$$y = \left(\frac{x - x_{\mu}}{x_{\sigma}} \right) \tag{2}$$

2.2.2   MIMO and MISO model development

The ANN consists of a number of inputs and outputs that are mapped together. The architecture of the model is a network between the input layer, hidden layer and output layer. Complex estimations were performed with the help of a nonlinear transfer function for the hidden layer whose features are controlled by the weight of the network. The learning activity is stopped after the prediction error falls below the specified error. A single layer FFNN is developed for both of the MIMO and MISO models.

After the training is complete, the model is validated using two different sets of validation data. This validation is performed to determine the optimum performance based on the average performance of the architectures. Next, the selected model is tested using independent testing data to confirm its performance. If a bad performance is obtained, the model must be retrained, and the steps are repeated until a good performance is observed (according to the

mean square error [MSE] and R^2). The detailed procedures are presented by Lei et al.[12] Finally, the selected optimum model is tested using the extrapolated data to determine its ability for predicting data outside of its training range.

2.2.3   ANN extrapolation capability test

Typically, the optimum ANN model is capable of predicting the outputs within the training range. Thus, the capabilities of the optimum MIMO and MISO models for predicting extrapolated data were tested. A set of extrapolation data were simulated based on several possible scenarios and a constant reflux ratio was applied. The simulation results produced x_b values ranging from 0.153 to 0.9996 kmol kmol^{-1} for the IPM in the reboiler. Meanwhile, the training data covered a range of 0.139 to 0.942 kmol kmol^{-1}. The extrapolation abilities of the MIMO and MISO were evaluated for this scenario because the x_b value exceeded the upper limit of the training range (0.942 kmol kmol^{-1}). An extrapolation test was performed to simulate the actual situation where a different operation was used. Thus, if the model can estimate out-of-range data with good agreement, the ANN model does not need the data for all operation ranges for training.

3.   RESULTS AND DISCUSSION

The training data were used to achieve the performance goal, and the ANN model architectures were stored with different weights and biases. Next, the developed MIMO and MISO models were validated with two sets of validation data, which resulted in 12 input nodes, 12 hidden nodes and 2 output nodes with [12-12-2] and [11-12-1] for the optimum MIMO and MISO models, respectively. The detailed validation and testing results for the models are presented in Lei et al.[12] Figure 1 and Figure 2 show the extrapolation capabilities of the MIMO and MISO models for predicting x_b, respectively.
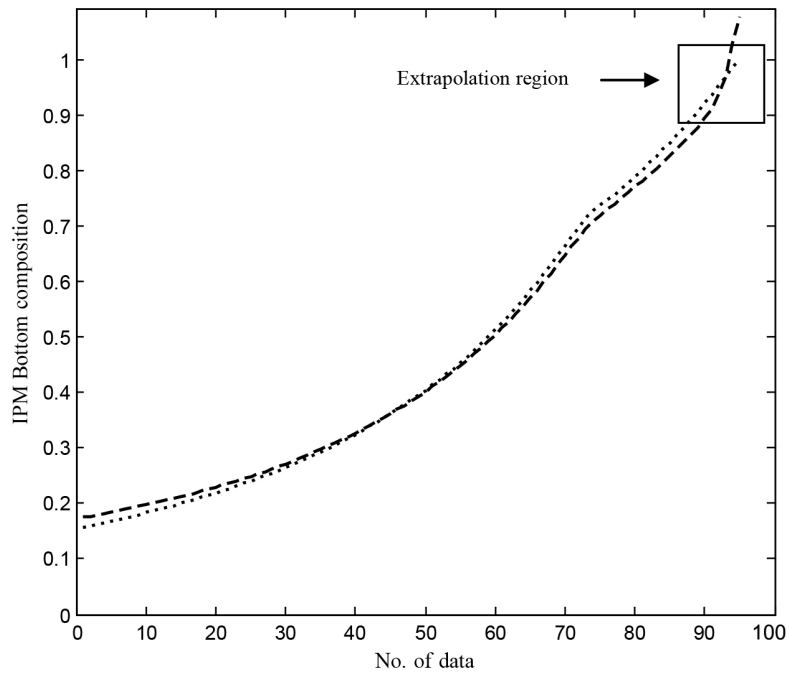
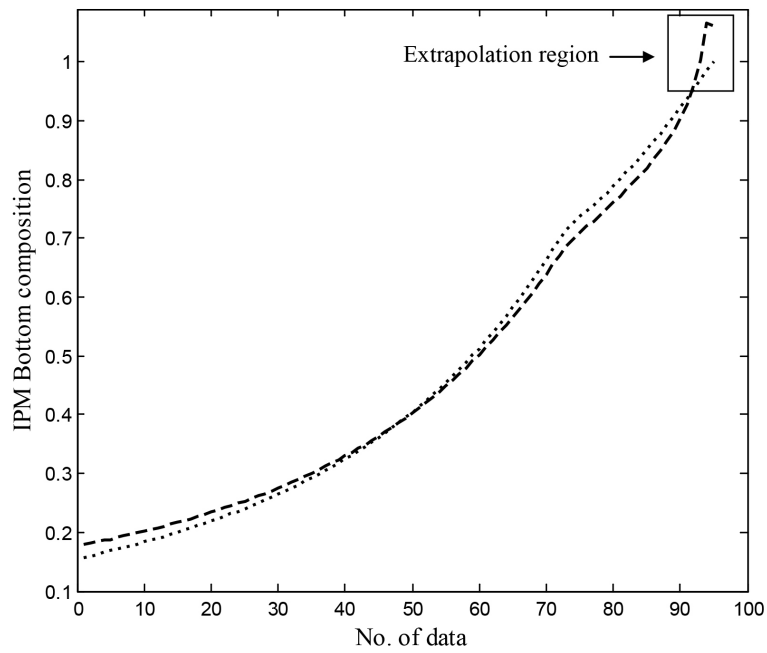Figure 1: Prediction of the MIMO model (--- NN ⋯ target).

Figure 2: Prediction of the MISO model (--- NN ⋯ target).

Figure 1 and Figure 2 show that the models satisfactorily predicted the x_b values with an MSE of 0.0078 and an R^2 of 0.9986 for the MIMO model and an MSE of 0.0063 and an R^2 of 0.9975 for the MISO model. When comparing these models, the MISO model resulted in better predictions than the MIMO model. This result occurred because the correlation between the input and output data of the MISO model was greater. Both figures show similar prediction trends with small deviations in the extrapolated data regions.

The failure of these models for predicting the extrapolated data well resulted from several factors. For example, insufficient historical data for catering to the dynamic process, the sensitivity of the normalisation technique to outliers, the nature of the nonlinear function inside the hidden neuron and the range of extrapolation data all contributed to the failure of the models. Most likely, the latter two processes control the robustness of the models for predicting extrapolated data. The out-of-training data will reach a constant value when it extends beyond the training range. Consequently, the ANN model cannot predict the data very well. In addition, this deviation results from a range of extrapolation data. Lei et al.[12] found that the increment of the extrapolation range can affect the prediction error. Nevertheless, satisfactory prediction was achieved. Although the results show that the models satisfactorily predicted the extrapolation data, the models can fit the interpolated data very well.

4. CONCLUSION

This paper examines the capability of the developed ANN MIMO [12-12-2] and MISO [11-12-1] models for predicting extrapolated bottom composition data. During the verification of the developed model on the testing data, both models predicted the data well with only small errors and little deviation when tested with the extrapolated data. The accuracies of the two models were compared, which indicated that the MISO model performed better because it produced a lower MSE value and a greater R^2 value than the MIMO model. To improve the robustness of the model, some modifications to the nonlinear transfer function must be recommended and the extrapolation limits must be identified.

5. ACKNOWLEDGEMENT

6.   REFERENCES

1.   Reusch, B. et al. (2001). *Neural networks as a tool for Gray box modelling in reactive distillation.* Berlin: Springer Berlin Heidelberg.
2.   Osorio, D. et al. (2004). Rigorous dynamic modeling and simulation of wine distillations. *Food Contr.*, 15(7), 515–521.
3.   Engell, S. & Fernholz, G. (2003). Control of a reactive separation process. *Chem. Eng. Process.*, 42(3), 201–210.
4.   Bahar, A. & Özgen, C. (2010). State estimation and inferential control for a reactive batch distillation column. *Eng. Appl. Art. Intel.*, 23(2), 262–270.
5.   Jithin Prakash, K., Patle, D. S. & Jana, A. K. (2011). Neuro-estimator based GMC control of a batch reactive distillation. *ISA Trans.*, 50(3), 357–363
6.   Konakom, K., Kittisupakorn, P. & Mujtaba, I. M. (2010). Neural network-based controller design of a batch reactive distillation column under uncertainty. *Asia-Pac. J. Chem. Eng.*, 7(3), 361–377.
7.   Mujtaba, I. M. & Greaves, M. A. (2006). Neural network based modeling and optimization in batch reactive distillation. IChemE symposium series 152, 868–876.
8.   Li, P. et al. (1998). Optimization of a semibatch distillation process with model validation on the industrial site. *Indust. Eng. Chem. Res.*, 37(4), 1341–1350.
9.   Seader, J. D. & Henley, E. J. (2006). *Separation process principles*. New York: John Wiley & Sons.
10.   Jimoh, M. et al. (1999). Transesterification of methyl myristate in a continuous reactive distillation column: Simulation and experiment. *Lipid/Fett.*, 101(2), 50–56.
11.   Bashah, N. A. A., Othman, M. R. & Aziz, N. (2013). Modeling of industrial scale semibatch reactive distillation using neural network. 6th International Conference on Process System Engineering (PSE Asia), Kuala Lumpur.
12.   Lei, Z. & Qi-Jun, Z. (2010). Simple and effective extrapolation technique for neural-based microwave modelling. *IEEE Micro. Wire. Comp. Lett.*, 20(6), 301–303.