CHEMICAL ENGINEERING TRANSACTIONS
VOL. 45, 2015
A publication of AIDIC
The Italian Association of Chemical Engineering
www.aidic.it/cet
Guest Editors: Petar Sabev Varbanov, Jiří Jaromír Klemeš, Sharifah Rafidah Wan Alwi, Jun Yow Yong, Xia Liu

ISBN 978-88-95608-36-5; **ISSN** 2283-9216
DOI: 10.3303/CET1545306

Feed Forward Neural Network Model for Production of Isopropyl Myristate in a Semibatch Reactive Distillation: An Evaluation of Extrapolation Prediction Capability

Nur Alwani Ali Bashah[a], Mohd Roslee Othman[b], Norashid Aziz*[,b]

[a]Facultyof Chemical Engineering, Universiti Teknologi Mara (Pulau Pinang), 13500 Permatang Pauh, Penang, Malaysia
[b]School of Chemical Engineering Campus, Universiti Sains Malaysia, Seri Ampangan, 14300 Nibong Tebal Penang, Malaysia
chnaziz@usm.my

Since last decade, the application of neural network (ANN) has been grown in chemical industries especially for the model based control system due to its capability to solve the complex model and its feasibility for online application. In general, the development of a good ANN model is depending on the quality of the data and the model structure. However, the ANN has a limitation in predicting extrapolation data. Thus in this paper, two multiple inputs - multiple outputs (MIMO) models (MIMO1 and MIMO2) were developed for production of isopropyl myristate in a semibatch reactive distillation and the capability of predicting the extrapolation data is evaluated by z-score normalization technique. Two data sets are simulated based on two possible scenarios occurs in industry: the first scenario is when the constant reflux ratio is applied for 10 h of batch time; the second scenario is when there is a high excess of isopropanol in the reboiler at the end of the process. The result shows that by using the z-score normalization, the MIMO2 was able to predict top composition (x_d) better than bottom composition (x_b) Since the MIMO2 model shows the ability to generalize extrapolation data for both X_b and X_d with low mean square error (MSE) and high coefficient of determination (R^2) value, thus it proves that using the z-score normalization method can facilitate the model to extrapolate the data satisfactorily.

1. Introduction

An Artificial Neural Network (ANN) is an empirical model that widely used to solve the complex task in various areas. Currently, the ANN model is adopted in the process system in order to overcome the disadvantages of fundamental model which is suffered on the large number of complex equations that expensive to be solved. ANN can also be used for modelling highly nonlinear processes. The structure of the ANN can be more complex and hence more representative than other empirical models and quite flexible model (Himmelblau, 2000). Batch reactive distillation (BRD) process is identified as highly nonlinear chemical process. Since it integrating both reactor and distillation column in a single unit that provides a complex interaction in term of thermodynamic and chemical equilibrium at every single stage. Thus, developing a theoretical rigorous model for the complex model like BRD process may not be practical for control and soft sensor application. As a result, the empirical model such as ANN can be introduced as the best option to model the BRD. ANN has been widely used to solve mathematical problems in BRD such as control system (Engell and Fernholz, 2003), soft sensor such as ANN state estimator and inferential control (Bahar and Özgen, 2010), neuro state estimator based generic model controller (Jithin, 2011) and optimization (Mujtaba and Greaves, 2006). By considering the advantages of the ANN model, it is an appropriate choice to embed the ANN model with control, soft sensor and optimization applications. In addition, the shorter computation time to evaluate problems makes the ANN is ideal to be implemented for real time optimization (Osuolale and Zhang, 2014). Usually, the product specification in the batch processes may vary from batch to batch and cause the output data sometime deviate from its original data boundary. It is known that the ANN model is mostly good in predicting data

Please cite this article as: Ali Bashah N.A., Othman M.R., Aziz N., 2015, Feed forward neural network model for production of isopropyl myristate in a semibatch reactive distillation: an evaluation of extrapolation prediction capability, Chemical Engineering Transactions, 45, 1831-1836 DOI:10.3303/CET1545306

between its trained data boundary but poorly estimate the out-of-range data i.e. extrapolation data. Thus, the developed ANN model needs to be robust so that the model is able to predict the output satisfactorily inclusive the out-of-training range data. In order to make the ANN model robust, i.e. insensitive to the presence of the outliers for the estimation of any necessary parameters, this problem needs to be catered. None of the available research on the ANN model of BRD process reported on the capability of the model to predict the extrapolation data. Prior to the model development, the process data need to be normalized so that the large magnitude gap between the parameters can be reduced. Thus, it can assure the equal influence on the different parameter in the model. Jayalakshmi and Santhakumaran (2011) had performed a study on the influence of the various types of normalization technique including z-score normalization technique to enhance the reliability of the trained network to diagnose the diabetes. Their result shows that, by adopting different normalization method, it can transform the model to be more robust. In this work, the ANN model based on z-score normalisation technique for BRD process is developed and its extrapolation capability is evaluated.

2. Neural network model development

2.1 Process simulation

The transesterification process of methyl myristate (MM) and isopropanol (IP) in an industrial scale semi-batch reactive distillation (BRD) was simulated using Batchfrac in Aspen Plus ®. All the specification in the BRD was based on the work conducted by Li et al. (1998) and the detailed process specification can be determined from Table 1.

Table 1: Process specification

Parameter	Specifications
Number of trays (including total condenser and reboiler)	32 trays
Condenser pressure	2.11 atm
Tray pressure drop	0.2072 atm
Column hold up	0.1 m^3
Distillate purity (methanol)	0.98 kmol/kmol
Initial charge to reboiler	72.64 kmol
Initial charge (methyl myristate, isopropanol, methanol))	24.57, 29.68, 18.39 kmol
Feed (isopropanol)	1.95 kmol

Prior to simulation, the column pressure was computed based on the procedure in Seader and Henley (2006). The transesterification reaction is take place in the reboiler and the reaction kinetics was taken from Jimoh et.al (1999).At the beginning of the process, the methyl myristate (MM), isopropanol (IP) and methnol (M) were charged into the reboiler together with the homogeneous catalyst. After 0.5 h of total reflux, the process was simulated for 26 h of batch time while the reaction and separation is taken place simultaneously. The lightest component (methanol) will be distillate while the main product isopropyl myristate (IPM) will be remained in the reboiler. By coupling reaction and separation like BRD, the reaction process have tendency to shift to the product side and the azeotropic condition can be suppressed during separation process. Hence, the productivity and yield are increased. All the process variables designed in the simulation were used to achieve 98 % of methanol (M) purity at distillate and complete conversion of MM in the reboiler. The detail of the simulation, the process operation and its specification can be found in Bashah et al. (2013). The developed Batchfrac model was used to generate the inputs-outputs data for ANN model development.

2.2 Data generation and Pre-processing data

In order to develop ANN model, the large number of inputs and outputs data needed were generated using validated Batchfrac model. However, only a few parameters which provide the significant impact to the process were selected based on work done by Bashah et al. (2012). Consequently, five significant input parameters were chosen i.e. reflux rate (R), reboiler heat duty (Q_b), feed flow rate (F), initial mol of isopropanol ($n_{IP,o}$) and methyl myristate ($n_{MM,o}$) in the reboiler. While the composition of methanol (x_d) and isopropyl myristate (x_b) at distillate and reboiler were chosen as output parameters. Besides that, the stage temperature (T) also observed to change significantly when the changes was made to any input parameters. Thus, it also included as the inputs parameter of the ANN model. Eight sets of different output patterns were simulated and divided into five training data sets, two validation data sets, and one testing data set. In addition, two sets of data were simulated in order to test on the extrapolation capability of the selected MIMO model. Another two data sets were simulated based on two possible scenarios occurs in

industry; the first scenario was when the constant reflux ratio was applied for along 10 h of batch time; the second scenario was when there was a high excess of isopropanol in the reboiler at the end of the process. All these data sets were normalized first before the ANN models were developed. The purpose of data scaling was to reduce the domination of the larger magnitude over the smaller ones. In this work, the z-score normalisation technique was applied to all the raw inputs and outputs in the data sets as shown in Eq(1). Where,y is normalized data, x is raw data, x_μ is mean data and x_σ is standard deviation data.

$$y = \left(\frac{x - x_\mu}{x_\sigma} \right) \tag{1}$$

2.3 Neural network architecture

The ANN is basically learning from experience where it learns the relationship between inputs and outputs data. The simple Feed Forward Neural Network (FFNN) model was developed which the architecture consists of input, hidden and output layers. The inputs parameter in the model include the real time (t) selected inputs from sensitivity analysis, real time stage temperature (T) and the historical data (t-1) from the inputs and outputs parameter as shown in Eq(2) and Eq(3).

MIMO1,

$$\left[x_d(t),\, x_b(t)\right] = f\left(R(t),\, Q_b(t),\, F(t),\, T(t),\, x_d(t-1),\, x_b(t-1),\, n_{IP,o},\, n_{MM,o}\right) \tag{2}$$

MIMO2,

$$\left[x_d(t),\, x_b(t)\right] = f\begin{pmatrix} R(t),\, Q_b(t),\, F(t),\, T(t),\, x_d(t-1),\, x_b(t-1),\, n_{IP,o},\, n_{MM,o}, \\ R(t-1),\, Q_b(t-1),\, F(t-1),\, T(t-1) \end{pmatrix} \tag{3}$$

The input layer was associated with the weight was summed and connected to the hidden node and transformed to the hidden node transfer function to produce the output. The number of node at input and output layers were depended to the number of the inputs and outputs of the model, while the hidden layer was based on the trial and error. The hyperbolic tangent sigmoid (Tan-Sig) transfer function was used for hidden layer while Purelin was used for output layer. The FFNN was trained by using Lavenberg-Marquardt (LM) learning algorithm and the training was stopped when it was reached desired performance goal of $1x10^{-4}$. The optimum architecture of ANN was determined based on the highest average of R^2 and lowest average of MSE towards the validation data. After the optimum MIMO model has been selected, the MIMO model was tested with the independent testing data. If the good generalization is obtained, then the MIMO models were further test with the extrapolation testing data sets. Otherwise, some modifications need to be done to the model architecture to improve their generalization ability.

3. Result and discussion

All the training data were trained to reach the desired performance goal and each of the models stored different random weights and biases. All these models were validated with the validation data to obtain the optimum MIMO model for both MIMO1 and MIMO2. MIMO1 model, which consists of 8 input, 11 hidden and 2 output nodes [8-11-2] and MIMO2, [12-12-2] were determined as optimum models to represent BRD process. Then, the optimum models were further tested with testing data to observe their generalization ability. Table 2 shows the MSE and R^2 value of validation and testing for optimum MIMO1 and MIMO2 models. From the result shown in Table 2, both models shows the good fit in the prediction of the X_b and X_d when tested with testing data.

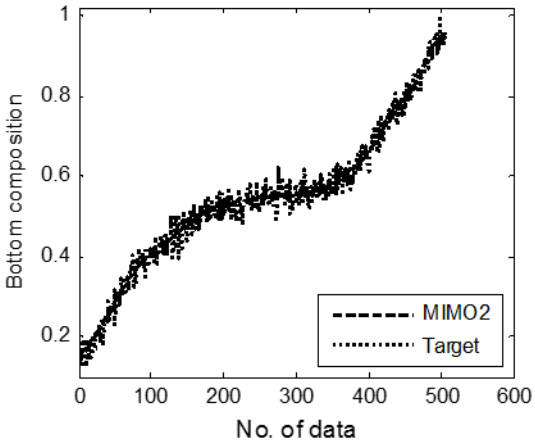
Table 2: MSE and R^2 value of Validation and Testing for the optimum MIMO1 and MIMO2 model

ANN Model	Architecture	Validation (MSE)	Validation, $x_d(R^2)$	Validation, $x_b(R^2)$	Testing (MSE)	Testing, $x_d(R^2)$	Testing, $x_b(R^2)$
MIMO1	[8-11-2]	0.0004	0.9997	1.0000	0.0010	0.9994	0.9998
MIMO2	[12-12-2]	0.0003	0.9998	1.0000	0.0006	0.9998	0.9999

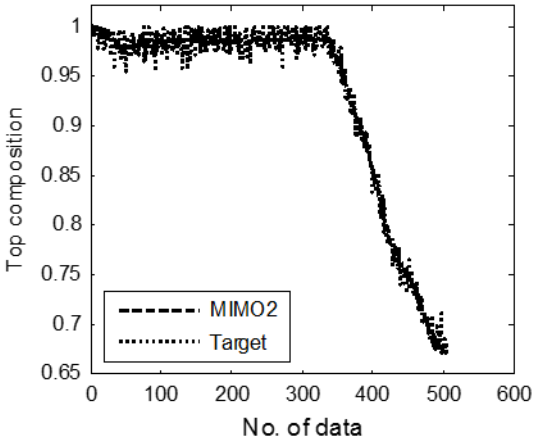
Although both models can generalize the validation data very well, but MIMO2 shows better MSE and R^2 than MIMO1. Thus, MIMO2 was chosen to represent BRD process for further test to fit the extrapolation

data. In order to mimic the similar condition in real industry, the random noise has been associated to the generated testing data to create the system complexity. Since the MIMO2 can predict X_b and X_d well, it was used to predict the system which was influenced by the noise. Figure 1 shows the profile predicted by MIMO2 model for both X_b and X_d. The MSE shows 0.0124 and R^2 0.9924 for X_b. While for $X_{d,}$, the MSE shows 0.9953 and R^2 is 0.9924. Although the prediction accuracy is reduced when noise is introduced, the model still shows its ability to follow the trend of the true testing data with small MSE value. Thus, it can be concluded that the developed MIMO2 model is robust which can yield good prediction although the data is contaminated with the noise data.

After proving the capability of MIMO2 to predict the data with and without noise, this model is used to evaluate its capability to predict the out of training range data i.e. extrapolation data. The first scenario used to evaluate the capability of the MIMO2 model to predict $X_{b,}$ while the second scenario is predicting X_d. Figure 2 shows the MIMO2 model satisfactorily predicted the bottom and top composition based on the specific case study.

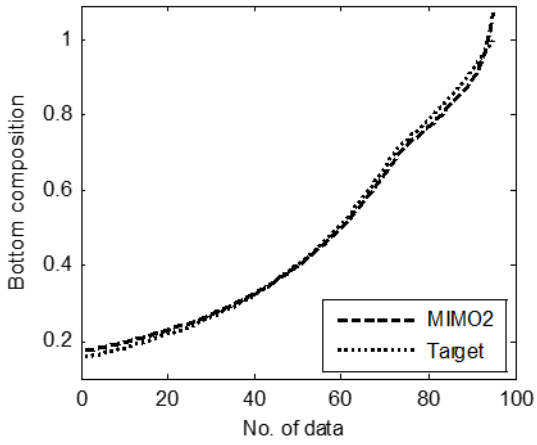


(a) Bottom Composition (X_b)

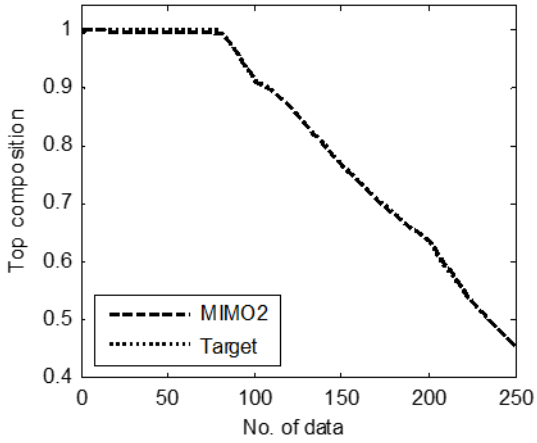


(b) Top composition (X_d)

Figure 1: Composition profiles with noise predicted by MIMO2 model



(a) Bottom Composition (X_b), Scenario 1



(b) Top composition (X_d), Scenario 2

Figure 2: Composition profiles predicted extrapolation data by MIMO2 model

In general, the model is able to predict well both extrapolation data with low MSE and high R^2 value. MIMO2 predicted X_b with MSE of 0.0078 and R^2 0.9986 while MSE for X_d is 0.0003 and R^2 is almost 1. The profiles show in Figure 2(a) and 2(b) also seems to follow the profiles of targeted value for both bottom and top composition. However, a small deviation is observed at the extrapolation data (data no. 92-95) as shown in Figure 2(a), while the MIMO2 line in Figure 2(b) shows good prediction on extrapolation data (data 243-250). The major different between these two scenarios is the percentage deviation of the extrapolation data from the trained boundary data. The extrapolation data showed 6 % higher for data 1 and 3.5 % lower for data 2 than the boundary limits of the trained data. In addition, the nature of the transfer functions used in the hidden and output layer can also contribute to the limitation of predicting the extrapolation data. Thus, determination of the acceptable range of the extrapolation data, especially for the dynamic process is important in order to ensure the model can predict BRD process until certain extends very well. It can be concluded that, the used of z-score normalization technique during pre-processing data and includes the historical inputs in the input layer can facilitate the model to extrapolate the data satisfactorily.

4. Conclusion

This work has demonstrated the potential of the MIMO model to predict the extrapolation data by implementing the z-score normalization technique in the pre-processing data. Two scenarios have been

simulated to evaluate the performance of the model to predict the bottom composition for scenario 1 and the top composition for scenario 2. After comparing the performance of MIMO1 and MIMO2 by cross-validation data, MIMO2 was chosen as the optimum model. The robustness of the model has also been tested by introducing with the noise on the BRD process data. Then it was further tested with the extrapolation data. Throughout the analysis, the MIMO2 model was able to generalize interpolation data with and without noise very well. The model was also able to predict the bottom and top composition of extrapolation data satisfactorily. Therefore, it shows that, implementing z-score normalization technique can enhance the reliability and robustness of the trained network. To further improve the robustness of the model, thus it is suggested to have some modifications on the nonlinear transfer function and the extrapolation limits must be properly identified.

Acknowledgement

The financial support from Universiti Teknologi Mara through training grant and Universiti Sains Malaysia through research university and incentive grant is gratefully acknowledged.

References

Bahar A., Özgen C., 2010, State estimation and inferential control for a reactive batch distillation column, Engineering Applications of Artificial Intelligence, 23**,** 262-270.

Bashah N.A.A., Othman M.R., Aziz N., 2012, Simulation of transesterification of methyl myristate and isopropanol in semibatch reactive distillation column, 4th AUN/SEED-Net Regional Conference on Chemical Engineering, 327-336 (in Malaysian).

Bashah N.A.A., Othman M.R., Aziz N., 2013, Neural network MIMO model for production of isopropyl myristate in a semibatch reactive distillation column, Applied Mechanics and Materials, 284, 403-408.

Engell S., Fernholz G., 2003, Control of a reactive separation process, Chemical Engineering and Processing, 42**,** 201-210.

Himmelblau D.M., 2000, Applications of artificial neural network in chemical engineering, Korean Journal of Chemical Engineering, 17**,** 373-392.

Jayalakshmi T., Santhakumaran A., 2011, Statistical normalization and back propagation for classification, International Journal of Computer Theory and Engineering, 3**,** 89-93.

Jimoh M., Garcia H.A., Wozny G., Bock H., Gutsche B., 1999, Transesterification of methyl myristate in a continuous reactive distillation column: Simulation and experiment, Lipid-Fett, 100(2), 50-56.

Jithin P.K., Patle D.S., Jana A.K., 2011, Neuro-estimator based GMC control of a batch reactive distillation, ISA Transactions, 50, 357-363.

Li P., Garcia H.A., Wozny G., Reuter E., 1998, Optimization of a semibatch distillation process with model validation on the industrial site, Industrial & Engineering Chemistry Research, 37**,** 1341-1350.

Mujtaba I.M., Greaves M.A., 2006, Neural network based modeling and optimization in batch reactive distillation, IChemE symposium series, 152, 868-876.

Osuolale F.N., Zhang J., 2014, Energy efficient control and optimization of distillation column using artificial neural network, Chemical Engineering Transactions, 39, 37-42,

Seader J.D., Henley E.J., 2006, Separation Process Principles, John Wiley & Sons (Asia), Pte Ltd., Singapore.