Optimization of artificial neural network topology for membrane bioreactor filtration using response surface methodology

Syahira Ibrahim, Norhaliza Abdul Wahab, Fatimah Sham Ismail, Yahaya Md Sam
School of Electrical Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia

Article Info	ABSTRACT
	The optimization of artificial neural networks (ANN) topology for predicting permeate flux of palm oil mill effluent (POME) in membrane bioreactor (MBR) filtration has been investigated using response surface methodology (RSM). A radial basis function neural network (RBFNN) model, trained by gradient descent with momentum (GDM) algorithms was developed to correlate output (permeate flux) to the four exogenous input variables (airflow rate, transmembrane pressure, permeate pump and aeration pump). A second-order polynomial model was developed from training results for natural log mean square error of 50 developed ANNs to generate 3D response surfaces. The optimum ANN topology had minimum ln MSE when the number of hidden neurons, spread, momentum coefficient, learning rate and number of epochs were 16, 1.4, 0.28, 0.3 and 1852, respectively. The MSE and regression coeffcient of the ANN model were determined as 0.0022 and 0.9906 for training, 0.0052 and 0.9839 for testing and 0.0217 and 0.9707 for validation data sets. These results confirmed that combining RSM and ANN was precise for predicting permeates flux of POME on MBR system. This development may have significant potential to improve model accuracy and reduce computational time.

Corresponding Author:

Norhaliza Abdul Wahab,
Control & Mechatronics Engineering, School of Electrical Engineering,
Faculty of Engineering,
Universiti Teknologi Malaysia, Johor, Malaysia.
Email: aliza@fke.utm.my

1. INTRODUCTION

Malaysia is one of the world's leading producers in palm oil industries [1]. Along with the increase of production capacity of palm oil every year, a large amount of wastewater was also being generated. These uncontrolled discharges of untreated palm oil mill effluent (POME) may cause pollution to the waterways [2]. In comparison with conventional activated system, membrane system is preferable to treat POME due to its simple operation, easy to scale-up, less weight and space requirements and high efficiency [3]. Membrane bioreactor (MBR) has been proven as a reliable technology in treating a wide range of water such as wastewater, groundwater and surface water. However, fouling phenomena is the main drawback of MBR system which contribute to high energy consumption and maintenance cost [4]. According to [5-7], fouling may varies with time during operation and this variation can be minimized by controlling the fouling variables [8].

Fouling can be controlled and reduced using several hydrodynamic condition techniques such as air bubble (aeration) control, relaxation, backwashing, and chemical cleaning [8-10]. It was found that still little

work conducted on the development of modelling and optimization for the operation condition of POME using MBR. Most of the works focused mainly on biological reduction of POME using MBR filtration [9-12]. Modeling of membrane process, involving with large number of parameters that needs to be considered is not an easy task.

Recently, modeling of membrane process using neural network has received enormous attention because of their ability in modeling and prediction of complex processes. ANN has been successfully applied to predict oily wastewater [13-14], permeate flux of albumin from serum bovine [15] and palm oil mill wastewater [16-17]. In addition, a good understanding of factors that affect ANNs model performance is crucial to predict the optimum value of the number of iterations, learning rate, momentum coefficient, number of hidden layers and number of hidden neurons. The parameters are varied until their optimal value are determined [18]. Determination of the best ANN topology is important because it affects the weight and bias. Usually it performed by trial and error [19-20] or one-variable-at-time (OVAT) [21-22] where this procedure is very time-consuming and monotonous task. According to [23] for three different level of each ANN variables, about 245 ($=3^5$) different configuration of ANN would be required. There is no specific rule used in selecting the value of variables in ANN. It is dependent on the complexity of the modeled system. Thus, it is of importance for researchers in order to find a standard technique to solve the problems associated with the ANN development.

Response surface methodology (RSM) as a collection of statistical and mathematical techniques has a capability for optimizing objective functions. It is a powerful optimum design tool in many engineering applications and can provide accurate models. RSM technique has been used to determine the ANN topology applied for multi-layer feed forward with backpropagation neural network [23-24]. It is also used to find the optimum value of neuron number in first and second hidden layers [18]. This paper aims for the development of radial basis function neural network (RBFNN) models for prediction of permeate flux during MBR filtration of POME wastewater. In this case, the RSM is proposed to find the optimum ANN topology to achieve minimum mean square error to improve the performance of the model

2. RESEARCH METHOD

2.1. Data collection

The experiments were carried out using membrane bioreactor for palm oil mill effluent (POME) with working volume of 20 L. The sample of POME was taken from Sedenak Palm Oil Mill Sdn. Bhd. in Johor, Malaysia with the working temperature at 27 ± 1 °C. There are four input variables for the POME model including transmembrane pressure (TMP), airflow rate, permeates pump and aeration pump. The output variable is permeate flux. The analysis of required data was carried out by using MATLAB R2014a and Design Expert version 7.1.6 to obtain the response surface and the contours plot. The total of 1602 data for each parameter were collected from the experiment including airflow rate, TMP, permeate pump, aeration pump and permeate flux. Figure 1 shows the flux was rapidly decreased after the airflow rate was decreased from 8 SLPM to 5 SLPM.
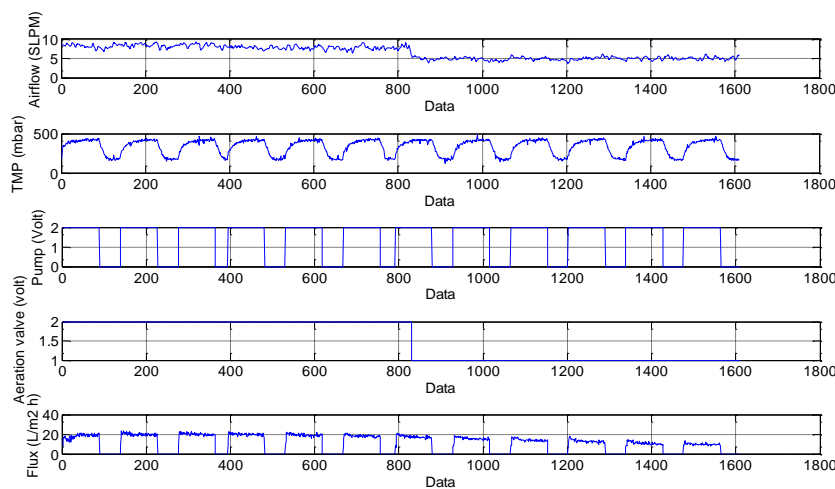

Figure 1. Data from MBR filtration experiment

2.2. Model development

In this work, the RBFNN model was used to predict the permeate flux of POME membrane bioreactor. Before that, all data need to undergo data pre-processing stage so called normalization. Since the input data for this system involved with different magnitude value and scale, all data were normalized into a minimum of +0 and maximum of +1. This procedure prevents the transfer function model from becoming saturated [25]. Equation (1) used for normalization given as:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

where X' is the scale value, X is the sample value while X_{min} and X_{max} are minimum and maximum value of data. The permeate flux was determined as given in (2):

$$J = \frac{V}{At} \tag{2}$$

where J is the permeate flux in $(lm^{-2}h^{-1})$, V is the volume flow rate in liter, A is membrane surface area (m^2) and t is the time (h). To investigate the feasibility of the predictive model, the collected data were separated into three data sets. From the total, 651 for training data set, where this data included the transition between high and low airflow rate. The 500 for testing data set was taken from the high airflow and finally, 451 for validation data set was taken from the low airflow rate. The training data was used to compute the network parameters. The testing data was used to assess the predictive ability of the generated model, while the remaining validation data was subsequenty used to ensure robustness of the network parameters and to avoid over-training [26]. The amount of training data set must be equal or larger than the amount of testing and validation data set to avoid extrapolation problem [27].

In this paper, three layers of RBFNN which are input, output and hidden were used. The non-linear transfer function of hyperbolic tangent sigmoid was used in the hidden layer and the linear transfer function of purelin was chosen for the output layer to produce a continuous output. The RSM is used to find the optimal value for each learning parameters of RBFNN model. 50 different experiments of central composite design (CCD) for five numerical factors (number of neurons, number of spread, learning rate, momentum rate and number of epoch) with eight repetition at center point were used. Five numerical factors and simulation ranks for RBFNN are shown in Table 1. The experimental results of the CCD were fitted with a second-order polynomial equation by a multiple regression technique. For predicting the optia point, the quadratic model is expressed by (3):

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i x_i + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \beta_{ij} x_i x_j + \varepsilon_{ij} \tag{3}$$

where Y is the response ln(MSE), $\beta_0, \beta_i, \beta_{ii}$ and β_{ij} are regression coefficients for intercept, linear quadratic and interaction coefficients, respectively and x_i and x_j are independent variables and k is a number of factors.

Table 1. The range of training parameters

RBFNN Parameter	Low	High
x_1: No. of neurons	1	20
x_2: Spread	0.1	2
x_3: Learning rate	0.01	0.4
x_4: Momentum rate	0.01	0.9
x_5: Number of Epoch	10	3000

All ANN topologies were designed and trained using RSM. The obtained quadratic equation was solved using response optimizer of RSM until the optimum condition to minimize MSE (response variable) data set was found. The MSE were transformed into natural log function (ln(MSE)) with α equal to 1. In this case, the distribution of the response variable become closer to the normal distribution [24].

2.3. Performance evaluation

$$MSE = \frac{1}{N} \sum_{i=1}^{n} (x_{pi} - x_{di})^2 \tag{4}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{n}\left(x_{pi} - x_{di}\right)^2} \qquad (5)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(x_{pi}-x_{di})^2}{\sum_{i=1}^{n}(x_{pi}-\bar{x})^2} \qquad (6)$$

where x_{pi} is the predicted output from observation i, x_{di} is the experimental or actual output form observation i, $\bar{x}$ is the average value of the experimental output and N is the number of data. Smaller values of MSE and RMSE mean a better performance of the model. For R^2 equal to 1 reveals that the regression line perfectly fit the data [26].

3.   RESULTS AND DISCUSSION

The relationship between the permeate flux and the independent parameters, namely number of neuron (x_1), spread (x_2), learning rate (x_3), momentum coefficient (x_4) and number of epoch (x_5) given as follows:

$$\ln(\text{MSE}) = -5.76 - 1.18\,x_1 - 0.68x_2 - 0.075x_3 - 0.041x_4 - 0.15x_5 + 0.18x_1x_2 + 0.077x_1x_3 + 0.043x_1x_4 + 0.15x_1x_5 - 0.051x_2x_3 - 0.043x_2x_4 - 0.12x_2x_5 - 0.043x_3x_4 + 0.040x_3x_5 + 0.041x_4x_5 + 0.81x_1^2 + 0.44x_2^2 + 0.010x_3^2 - 0.024x_4^2 + 0.081x_5^2 \qquad (7)$$

The fitness of the model is determined by analysis of variance (ANOVA) which consists of sum of square (SS), degree of freedom (df), mean square (MS), F-values and P-values as shown in Table 2. The significance of each coefficient was determined by the F-test and P-value. The significant of corresponding variables would be increase if the absolute F-value becomes greater and the P-value becomes smaller. From Table 2, the model gives F-value of 81.25 and very low P-value (< 0.0001). P-values < 0.05 reveal that the model terms were significant. The number of neuron had the highest effect on ln(MSE) response followed by number of spread and number of epoch. The learning rate and momentum coefficient had no significant effect on the responses. The prediction R^2 of 0.9825 is in reasonable agreement with adjusted R^2, 0.9704. The low value of coefficient of variance (CV=4.62%) which is less than 10 showed that the experiments conducted were precise and reliable.

Table 2. ANOVA for predicted RSM model

Source	SS	df	MS	F-value	P-Value	Prob > F
Model	84.45	20	4.22	81.25	< 0.0001	Significant
x_1 -Number of neuron	44.62	1	44.62	858.64	< 0.0001	
x_2 -Spread	14.74	1	14.74	283.53	< 0.0001	
x_3 -Learning rate	0.18	1	0.18	3.50	0.0714	
x_4 -Momentum coefficent	0.055	1	0.055	1.05	0.3137	
x_5 -Number of epoch	0.69	1	0.69	13.21	0.0011	
Residual	1.51	29	0.052			
Lack of Fit	1.51	21	0.072			
Pure Error	0.000	8	0.000			
Cor Total	85.96	49				
Model statistics						
Std. Dev.	0.23	R-Squared	0.9825			
Mean	-4.93	Adj R-Squared	0.9704			
C.V. %	4.62	Pred R-Squared	0.9261			

3.1. Response surface plot results

The plot of response surface results is presented in Figure 2. Each graph represented a combination of two factors at the time and holding all other factors at the middle level. Figure 2(a) shows the response surface ln(MSE) versus the number of neuron and spread while other factors remained constant at zero level. It can be seen from Figure 2(a) that minimum value of ln(MSE) can be found by 15-20 neurons and 1.5-2.0 spread. Moreover, this range was observed for neuron number in relation with epoch number as shown in Figure 2(b). The response surface show that along with an increase in number of epoch from 10 to 3000 and spread from 0.1 to 2.0, the ln(MSE) decreased to - 6.2 as shown in Figure 2(c). The objective function value

for ln(MSE) is -6.342 for the final points as presented in Figure 2(a). The optimum values given by RSM were as follows: number of neurons = 16, spread = 1.4, learning rate = 0.28, momentum rate = 0.3 and number of epochs = 1852. These optimum values are used in the training of RBFNN.
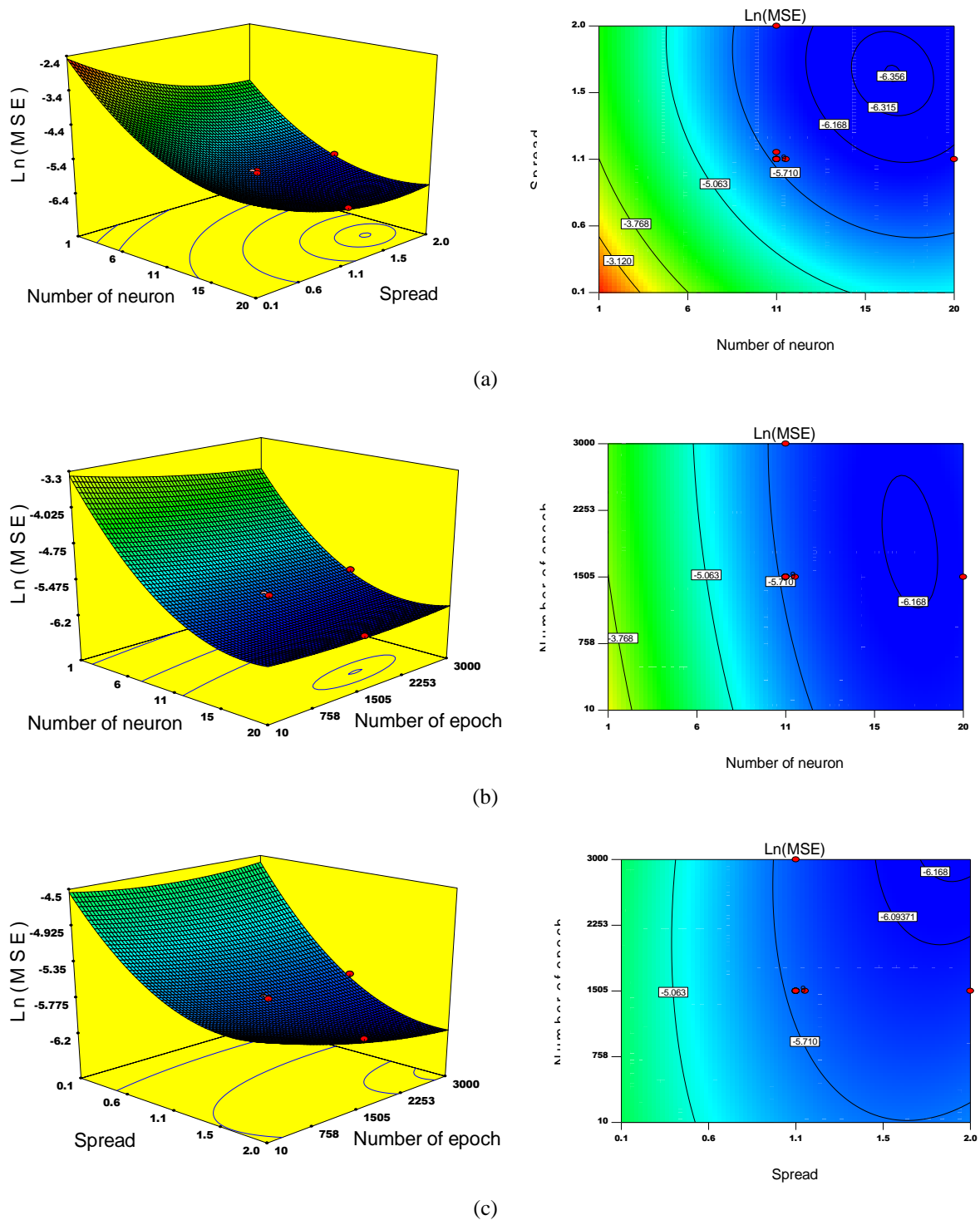

(a)

(b)

(c)

Figure 2. Response surface and contour plot of ln (MSE) for (a) neuron number and spread, (b) neuron number and epoch number, and (c) spread and epoch number while other factors remained constant.

3.2. Neural network plot results

In this section, the regression plots of the experimental data versus the computed neural network data using the optimum ANN topology are presented for each step incuding training, testing and validation netwoks. The predicted models were well fitted to the experimental data for all steps as depicted in Figure 3. The correlation coefficients (R) for training is 0.9906, for testing is 0.9839 and for validation is 0.9707. The comparative values correlation of determination (R^2), RMSE and MSE were given in Table 3. The results showed that the optimum ANN model is suitable for describing the permeate flux of POME using MBR filtration. The optimal topology of ANN using RSM provided good quality prediction for the five exogenous outputs. The results have been compared with the conventional RBFNN and showed an improved ANN model performance as shown in Table 3. The RBFNN-RSM showed its superiority and faster then trial-and-error methods in finding the optimum topology of ANNs.
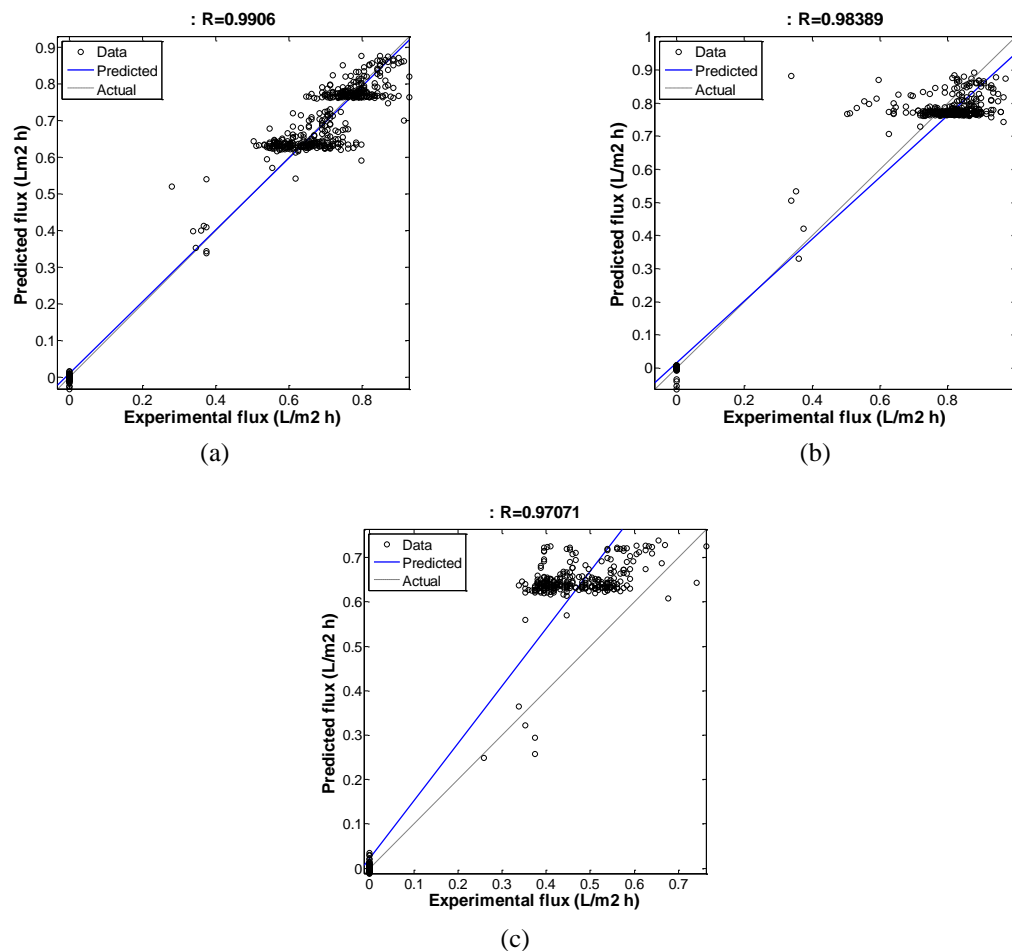

(a)

(b)

(c)

Figure 3. Regression plot for predicted versus experimental flux for (a) training data, (b) testing data and (c) validation data

Table 3. Performance evaluation

RBFNN-RSM	R^2	MSE	RMSE
Training	0.9813	0.0022	0.0470
Testing	0.9681	0.0052	0.0722
Validation	0.9423	0.0217	0.1473
Conventional RBFNN	R^2	MSE	RMSE
Training	0.9422	0.0076	0.0872
Testing	0.9374	0.0096	0.0980
Validation	0.7956	0.0200	0.1414

Figures 4(a)-(c) show the response variable of permeate flux for training, testing and validation, respectively. For training data which is the transition between high to low airflow rate, the permeate flux starts to decrease slowly from 0.88 to 0.60 L/m2 h. For testing data, the permeate flux is at high airflow rate and it remains at 0.8 L/m2 h. For validation data, the permeate flux decreases rapidly compared to the permeate flux at high airflow. It can be seen that good prediction models are obtained for the permeate flux for all data set.
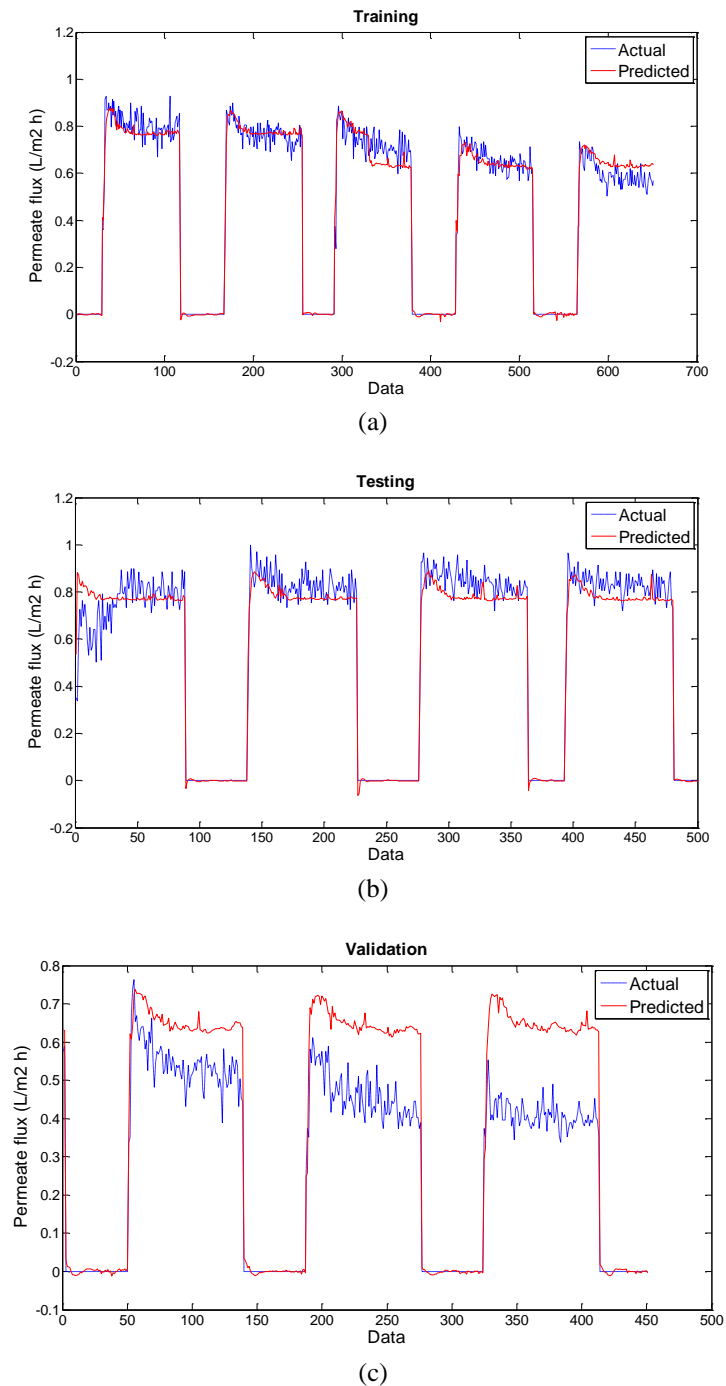

Figure 4. POME Permeate flux for (a) training data, (b) testing data and (c) validation data set.

4. CONCLUSION

The folowing conclusion can be drawn from the investigations conducted in this work: The optimal ANN topology of RBFNN topology was more precise for predicting permeate flux of POME using MBR with low MSE (0.0022) and high correlation coefficient (0.9906). The optimal neural model had minimum when the number of neurons, spread, learning rate, momentume rate and number of epochs runs were equal to 16, 1.4, 0.28, 0.3 and 1852, respectively. The results of testing and validation model on new trials showed excellent agreement between the actual and predicted data with correlation coefficient equal to 0.9839 and 0.9707, respectively. The application of integrated RBFNN and RSM reduces the computational cost and improved the ANN model prediction.

REFERENCES

[1] M. F. Alkhatib, A. A. Mamun, and I. Akbar, "Application of response surface methodology (RSM) for optimization of color removal from POME by granular activated carbon," *Int. J. Environ. Sci. Technol.*, vol. 12, pp. 1295–1302, 2015.

[2] W. P. Wah, S. Nik Meriam, M. Nachiappan, and B. Varadaraj, "Pre-treatment and membrane ultrafiltration using treated palm oil mill effluent (POME)," *Songklanakarin J. Sci. Technol.*, vol. 24, pp. 891–898, 2002.

[3] A. Cassano and A. Basile, "Membranes for industrial microfiltration and ultrafiltration," in *Advanced Membrane Science and Technology for Sustainable Energy and Environmental Applications*, 1st ed., A. Basile and S. P. Nunes, Eds. Elsevier, 2011, pp. 647–679.

[4] G. Mohd Syahmi Hafizi, T. Y. Haan, A. W. Lun, N. Abdul Wahab, Mohammad Rahmat, and M. Khairul Muis, "Fouling assessment of tertiary palm oil mill effluent (POME) membrane treatment for water reclamation," *J. Water Reuse Desalin.*, vol. 8, no. 3, pp. 412–423, 2018.

[5] T. Leiknes, "Wastewater Treatment by Membrane Bioreactors," in *Membrane Operations: Innovative Separations and Transformations*, E. Drioli and L. Giorno, Eds. Italy: WILEY-VCH Verlag GmbH & Co. KGaA, 2009, pp. 374–391.

[6] H. Lin *et al.*, "Membrane bioreactors for industrial wastewater treatment: A critical review," *Crit. Rev. Environ. Sci. Technol.*, vol. 42, no. 7, pp. 677–740, 2012.

[7] T. Janus, "Modelling and Simulation of Membrane Bioreactors for Wastewater Treatment," De Montfort University, Leicester, 2013.

[8] S. Judd, "Fouling control in submerged membrane bioreactors," *Water Sci. Technol.*, vol. 51, no. 6–7, pp. 27–34, 2005.

[9] N. H. Abdurahman, H. N. Azhari, and S. Nurdin, "An Integrated Ultrasonic Membrane An aerobic System (IUMAS) for Palm Oil Mill Effluent (POME) Treatment," *Energy Procedia*, vol. 138, pp. 1017–1022, 2017.

[10] Z. Ahmad, M. B. Ridzuan, and Z. Daud, "Membrane Bioreactor for Palm Oil Mill Effluent and Resource Recovery," in *International Conference on Sustainable Development for Water and Waste Water Treatment. December 2009*, 2009, pp. 1–8.

[11] S. Muhammad, M. Abdul Wahab, M. N. Mohd Tusirin, S. A. Siti Rozaimah, and A. H. Hassimi, "Investigation of Three Pre-treatment Methods Prior to Nanofiltration Membrane for Palm Oil Mill Effluent Treatment," *Sains Malaysiana*, vol. 44, no. 3, pp. 421–427, 2015.

[12] N. S. Azmi and K. F. M. Yunos, "Wastewater Treatment of Palm Oil Mill Effluent (POME) by Ultrafiltration Membrane Separation Technique Coupled with Adsorption Treatment as Pre-treatment," *Agric. Agric. Sci. Procedia*, vol. 2, pp. 257–264, 2014.

[13] R. Badrnezhad and B. Mirza, "Modeling and optimization of cross-flow ultrafiltration using hybrid neural network-genetic algorithm approach," *J. Ind. Eng. Chem.*, vol. 20, no. 2, pp. 528–543, 2014.

[14] R. Soleimani, N. A. Shoushtari, B. Mirza, and A. Salahi, "Experimental investigation, modeling and optimization of membrane separation using artificial neural network and multi-objective optimization using genetic algorithm," *Chem. Eng. Res. Des.*, vol. 91, no. 5, pp. 883–903, 2013.

[15] S. Curcio, V. Calabrò, and G. Iorio, "Reduction and control of flux decline in cross-flow membrane processes modeled by artificial neural networks and hybrid systems," *Desalination*, vol. 236, no. 1–3, pp. 234–243, 2009.

[16] Y. Zakariah, A. W. Norhaliza, S. Shafishuhaiza, and A. H. A. Raof, "Permeate Flux Measurement and Prediction of Submerged Membrane Bioreactor Filtration Process Using Intelligent Techniques," *J. Teknol. UTM*, vol. 73, no. 3, pp. 85–90, 2015.

[17] Y. Zakariah, A. W. Norhaliza, and S. Sahlan, "Modeling of submerged membrane bioreactor filtration process using NARX-ANFIS model," in *2015 10th Asian Control Conference: Emerging Control Techniques for a Sustainable World, ASCC 2015*, 2015, pp. 1–6.

[18] E. Razmi-Rad, B. Ghanbarzadeh, S. M. Mousavi, Z. Emam-Djomeh, and J. Khazaei, "Prediction of rheological properties of Iranian bread dough from chemical composition of wheat flour by using artificial neural networks," *J. Food Eng.*, vol. 81, no. 4, pp. 728–734, 2007.
[19] K. C. Keong, M. Mustafa, A. J. Mohammad, M. H. Sulaiman, and N. R. H. Abdullah, "Artificial neural network flood prediction for sungai isap residence," in *2016 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)*, 2016, pp. 236–241.
[20] D. Krishna and R. P. Sree, "Artificial Neural Network and Response Surface Methodology Approach for Modeling and Optimization of Chromium (VI) Adsorption from Waste Water using Ragi Husk Powder," *Indian Chem. Eng.*, vol. 55, no. 3, pp. 200–222, 2013.
[21] F. Ibrahim, M. N. Taib, W. A. B. Wan Abas, C. C. Guan, and S. Sulaiman, "A novel dengue fever (DF) and dengue haemorrhagic fever (DHF) analysis using artificial neural network (ANN)," *Comput. Methods Programs Biomed.*, vol. 79, no. 3, pp. 273–281, 2005.
[22] M. Mustafa, T. Mohd Nasir, H. M. Zunairah, S. Norizam, and M. A. Siti Armiza, "Classification of EEG Spectrogram Image with ANN approach for Brainwave Balancing Application," *Int. J. Simul. Syst. Sci. Technol.*, vol. 12, pp. 29–34, 2011.
[23] M. Aghbashlo, M. H. Kianmehr, T. Nazghelichi, and S. Rafiee, "Optimization of an artificial neural network topology for predicting drying kinetics of carrot cubes using combined response surface and genetic algorithm," *Dry. Technol.*, vol. 29, no. 7, pp. 770–779, 2011.
[24] T. Nazghelichi, M. Aghbashlo, and M. H. Kianmehr, "Optimization of an artifial neural network topology using couple response surface meethodology and genetic algorithm for fluidize bed drying," *Comput. Electron. Agric.*, vol. 75, pp. 84–91, 2011.
[25] I. Syahira, "Amylose content calibration model for the three types of selected rice grains using visible shortwave near infrared spectroscopy," Universiti Teknologi Malaysia, 2015.
[26] H. Nourbakhsh, Z. Emam-Djomeh, M. Omid, H. Mirsaeedghazi, and S. Moini, "Prediction of red plum juice permeate flux during membrane processing with ANN optimized using RSM," *Comput. Electron. Agric.*, vol. 102, pp. 1–9, 2014.
[27] K. Chia, A. R. Herlina, and A. R. Ruzairi, "Neural network and principal component regression in non-destructive soluble solids content assessment: a comparison.," *J. Zhejiang Univ. Sci. B*, vol. 13, no. 2, pp. 145–51, Feb. 2012.

BIOGRAPHIES OF AUTHORS

Syahira Ibrahim received her B. Eng. Hons (Electric-Control and Instrumentation) and MSc. of Engineering (Electrical) from Universiti Teknologi Malaysia in 2011 and 2015 respectively. She is currently working toward her PhD in process control at School of Elecrical Engineering, Universiti Teknologi Malaysia. Her current research interests include modeling of near-infrared spectroscopy and optimization of membrane filtration system using hybrid response surface methodology-artificial intelligent.

Ir. Dr Norhaliza Abdul Wahab is currently an Associate Professor at Universiti Teknologi Malaysia (UTM). She is currently the Director for Control and Mechatronic Engineering at the School of Electrical Engineering, UTM. She completed her PhD in Electrical Engineering majoring in Control in July 2009. She is actively involved in researching and teaching in the field of industrial process control. Her expertise is in modelling and control of industrial process plant. Recently she has worked primarily on different types of domestic and industrial water and wastewater treatment technology towards optimization and energy saving system.

Fatimah Sham Ismail is currently a Senior Lecturer at Universiti Teknologi Malaysia (UTM), Johor Bahru. She has more than 20 years experience in area of Control and Instrumentation Engineeringsince joining UTM in 1992. She received the B.Sc (Hons.) in Physics, 1989 from Universiti Kebangsaan Malaysia and obtained her Master and Ph.D from University Teknologi Malaysia (UTM) in 1992 and 2011, respectively. Currently, she is conducting researches on development optimization algorithm for multi-objective problems, plant optimization design, and fault detection & diagnosis.

Yahaya Md. Sam received the B.E. degree in electrical engineering from University Technology of Malaysia in 1986, M.Sc. degree in control systems engineering from Sheffield University, United Kingdom, in 1988, and the Ph.D. degree in control engineering from University Technology of Malaysia in 2004. He is currently a Professor with the School of Electrical Engineering, University Technology of Malaysia. His research interests include an optimal control, robust control, composite nonlinear feedback and sliding mode control and application of these ideas to the automotive systems.