

PAPER • OPEN ACCESS

Modeling of vanillin adsorption from aqueous solution using resin H103 by artificial neural network

To cite this article: W S Chan *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **702** 012048

View the [article online](#) for updates and enhancements.

You may also like

- [Electrochemical Polymer and Polymer Additive Synthesis from Biobased Vanillin and Syringaldehyde](#)
Robin Kunkel, Dominik Müller, Detlef Schmiedl et al.
- [Effect of Vanillin to Prevent the Dendrite Growth of Zn in Zinc-Based Secondary Batteries](#)
Mukkannan Azhagurajan, Akiyoshi Nakata, Hajime Arai et al.
- [Adhesion of *Candida albicans* to Vanillin Incorporated Self-Curing Orthodontic PMMA Resin.](#)
K Zam, P Sawaengkit, S Thaweboon et al.



ECS The Electrochemical Society
Advancing solid state & electrochemical science & technology

242nd ECS Meeting

Oct 9 – 13, 2022 • Atlanta, GA, US

Presenting more than 2,400 technical abstracts in 50 symposia

Register now!

ECS Plenary Lecture featuring M. Stanley Whittingham,
Binghamton University
Nobel Laureate –
2019 Nobel Prize in Chemistry

The banner features the ECS logo, the meeting title and dates, a portrait of M. Stanley Whittingham with his Nobel Prize medal, and a 'Register now!' button. The background includes images of a conference audience and a person interacting with a futuristic interface.

Modeling of vanillin adsorption from aqueous solution using resin H103 by artificial neural network

W S Chan¹, R A Samah^{1,*}, N Zainol¹, A S Fakharudin², S A Aziz³ and L Y Phang³

¹Faculty of Chemical and Process Engineering Technology, College of Engineering Technology, Universiti Malaysia Pahang, Lebuhraya Tun Razak, 26300 Gambang, Kuantan, Pahang, Malaysia.

²Faculty of Computing, Universiti Malaysia Pahang, Lebuhraya Tun Razak, 26300 Gambang, Kuantan, Pahang, Malaysia.

³Department of Bioprocess Technology, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.

*E-mail: rozaimi@ump.edu.my

Abstract. Vanillin adsorption onto resin H103 was modelled using artificial neural network (ANN) approach and the best ANN algorithm was determined in this work. The first step of ANN modeling was ANN set up, followed by the optimization of ANN. The parameters for the input layers are contact time, initial vanillin concentration, resin dosage, pH, and temperature while the response is residual vanillin concentration. The neural network was trained using backpropagation (BP) algorithm. The result shows that the Levenberg–Marquardt algorithm was best suited the training function and the optimized ANN involved seven neurons at the hidden layer. This model can produce a correlation of determination value of 0.9999 with the mean square error (MSE) value of 0.0277. The best adsorption efficiencies for each factor were 98.11%, 96.03%, 98.14%, 98.2%, and 98.10% at 2.0 g of adsorbent dosage, 30 min of contact time, 100 mg/L of initial vanillin concentration, pH 5, and 25 °C, respectively. The outcomes of this work proved that ANN is excellent in predicting experimental data of vanillin adsorption by resin H103.

1. Introduction

Vanillin is one of the most important flavoring materials which has been widely used in the food industry, perfumes, cosmetics, and pharmaceuticals [1]. Vanillin can be recovered using several techniques such as distillation, extraction using gas-expanded liquid, ultrafiltration, and adsorption. The adsorption process is preferable to separate vanillin from lignin due to its nature of environmentally friendly and more economical compared to the other separation technologies [2].

The world annual production of vanillin is more than 12,000 tons. However, less than 1% of the material is natural vanillin originated from botanical sources such as cured vanilla pods and *Java citronella* [3]. The price of synthetic vanillin is less than USD15/kg, in contrast to the price of natural vanilla between USD1,200.00/kg and USD4,000.00/kg [4]. Because of the limited supply and high price of natural vanillin combined with the increasing demand for natural flavors, it becomes a motivation for the researchers to find an alternative method to produce vanillin.

Modeling is done before the implementation of a new method to predict its feasibility and explain the possibilities of response in particular circumstances. In the past few decades, it is challenging to do



modeling and optimization of adsorption because they have to carry out many experiments to obtain the desired and optimum value. In the 50s, response surface methodology (RSM) was proposed by Box and Behnken [5] for the purpose modeling and optimization, and this helps the researcher save a lot of time. However, modeling and optimization of the heterogeneous photon-Fenton process using RSM showed that the degradation efficiency of the experiment was within the range of 70% to 90%, not high enough to describe the behavior of the experiment [6]. The problem faced by the same data can be solved by using ANN due to its ability to achieve a degradation efficiency of 98.11% [7]. The improvement occurred because ANN consists of three layers and a training algorithm to adjust the error and repeat until the desired output is obtained [8].

ANN is a computing system which imitates the biological neural network of living organisms. It is based on a collection of connected units called artificial neurons. These artificial neurons are typically organized in three groups – input, hidden, and output layers. Activities of the hidden layers depend on the activities of the input layers, while the hidden layers will affect the output layer. The results are then processed by an input–output function (transfer function) [9]. The number of hidden layers is directly proportional to the accuracy of the result. Other than predicting the output variables, it is also used to eliminate irrelevant input, called pruning method. When the significant variable is determined, other irrelevant variables are eliminated from the model which can help to reduce the size of the network and minimize the redundancy in the training data [10].

This work aimed to model the vanillin adsorption from aqueous solution using resin H103 by ANN, and the results describe that a good agreement between the experimental data and predicted output was achieved. Resin H103 was selected due to its high adsorption capacity and low cost [11]. The effects of adsorbent dosage, contact time, vanillin initial concentration, pH and temperature were also studied [11]. The designation of a neural network, selection of the best training method, and identification of the suitability of ANN for modeling of vanillin adsorption played an essential role in completing this research.

2. Methodology

2.1. Data Preparation

The experimental data of vanillin adsorption from aqueous solution using resin H103 were taken from a previous study [11]. The experimental data show the effect of adsorbent dosage, contact time, initial vanillin concentration, pH, and temperature on the adsorption performance of resin H103. The selection of resin H103 was based on its highest capacity and low cost. For this modeling, the data of model variables are shown in table 1.

Table 1. Data used for ANN modeling [11].

Variables	Range
Input layer	
Adsorbent dosage (g)	0.5–5.0
Contact time (min)	3–300
Initial vanillin concentration (mg/L)	6.25–800
pH	3.00–7.00
Temperature (°C)	25–55
Output layer	
Vanillin residual concentration (mg/L)	0.41–44.92

2.2. Artificial neural network model

2.2.1. Backpropagation algorithm. Backpropagation (BP) was used as the training algorithm for the modeling of vanillin adsorption from aqueous solution using resin H103 [12]. It was used to change the

weights and biases in a network to improve the overall output of the network. The working principle of the backpropagation algorithm is iteration. The iteration will only stop when the error between the predicted output and the experimental data is acceptable [13]. There are 13 backpropagation algorithms exist in the ANN, and all were performed to determine the best suited BP training algorithm for this model. These training algorithms were performed with an initial ANN design of three layers with tangent sigmoid transfer function (TANSIG) at the hidden layer and linear transfer function (PURELIN) at the output layer. All the data were normalized to be between 0 and 1 to suit the application of sigmoid function in the ANN [14]. A linear activation function (PURELIN) is normally used for the output layer when the ANN is trained for regression before optimization. Ten neurons were used at this stage to determine the best-suited training algorithm. During the training, the experimental data were randomly divided into training, validation, and test subsets at a ratio of 2:1:1 [15]. The validation set was used to determine when to stop the iteration of the training set to prevent overfitting [16]. The iteration of the BP algorithm was stopped when the MSE of the validation set reached the minimum value and it started to bounce back.

In the next stage, the MSE for each training algorithm was compared. The performance of the BP training algorithm was measured using the MSE value. A smaller MSE value indicates that the predicted output is very close to the experimental data. Therefore, the BP training algorithm with the minimum value of MSE was selected as the best-suited training algorithm and to be used in the subsequent modeling.

2.2.2. Optimization of ANN. The best-suited training algorithm determined from the previous part was used in the three layers ANN with tangent sigmoid transfer function at the hidden layer and linear transfer function at the output layer to perform the optimization of ANN. The optimization of ANN was performed as the function of the number of neurons in the hidden layer and was evaluated based on the MSE and linear regression [10]. Two neurons were used as an initial guess to evaluate the performance of the ANN then it was repeated by increasing the number of neurons with an interval of one. The performance of ANN is improved with the increase in the number of neurons, but when the number of neurons exceeds the threshold number, overfitting of the model will occur and causes the model to fail to predict the experimental data [17]. Therefore, the number of neurons is increased until the MSE of the model started to increase after the minimum point was achieved.

Then, the MSE of the model for each number of neurons was compared. The optimum number of neurons was determined with the minimum value of MSE and the highest linear regression of the model. Finally, the three-layer ANN with tangent sigmoid transfer function at the hidden layer and linear transfer function at the output layer was evaluated at the best-suited training algorithm and the optimum number of neurons.

2.2.3. Validation of effect of variables on adsorption efficiency. The effects of adsorbent dosage, contact time, initial vanillin concentration, pH, and temperature on the experimental output were studied and validated by the predicted output from the ANN model. The output from the experimental data is residual vanillin concentration, which is correlated to the adsorption efficiency for the ease of comparison. The adsorption efficiency was calculated using equation (1):

$$\text{Adsorption efficiency (\%)} = \frac{100(C_1 - C_2)}{C_1} \quad (1)$$

where C_1 and C_2 are the initial and final vanillin concentrations in the solution, respectively. The adsorption efficiency from the experimental data was compared with the predicted output from the ANN. The predicted output which is deviated from the experimental data is justified. Finally, the suitability of ANN for the modeling of vanillin adsorption from aqueous solution using resin H103 was evaluated based on the regression of the model.

3. Results and Discussion

3.1. Selection of backpropagation (BP) algorithm

Thirteen BP algorithms were compared to determine the best training function to be used in the modeling. A three-layer of ANN with a tangent sigmoid transfer function (TANSIG) at the hidden layer and a linear transfer function (PURELIN) at the output layer was used to perform all the training algorithms. The hidden layer consisted of 10 number of neurons. Gradient descent with momentum learning function (LEARNGDM) was used for this network and mean squared error (MSE) was used to measure the performance of the network. The result shows that the Levenberg–Marquardt backpropagation algorithm (LMA) with a minimum MSE was the best training function for this modeling.

The comparison of the different algorithms was based on the deviation of the predicted output from the experimental data. A large deviation of some algorithms from the experimental data was observed, while some of the algorithms were able to converge to the experimental data. The results show that LMA was able to produce the smallest deviation compared to the other backpropagation algorithms. In contrast, the predicted output by gradient descent algorithm (GDA) has the largest deviation. As shown in table 2, a minimum MSE with a value of 0.024 was produced by TRAINLM, which is followed by TRAINBR with the MSE value of 0.114. The MSE values produced by other training algorithms such as TRAINBFG, TRAINCGB, TRAINCGF, TRAINCGD, TRAINGDM, TRAINGDA, TRAINGDX, TRAINOSS, TRAINRD, and TRAINSCG are greater than one. The largest MSE was produced by TRAINGD with the MSE value of 1500. The deviation produced by the algorithm can be attributed to the nonlinear structure and combinatorial nature of the experimental data. Therefore, LMA is the best training function for this modeling because it can solve nonlinear least-squares problems and produce a least-squares curve fitting [18].

Table 2. Comparison of 13 backpropagation algorithms.

Backpropagation (BP) algorithms	Function	Mean square error (MSE)	Iteration number (IN)	R^2	Best linear equation (BLE)
BFGS quasi-Newton	TRAINBFG	9.910	23	0.919	$y = 0.92x - 0.055$
Bayesian regularization	TRAINBR	0.114	1000	1.000	$y = x + 0.0075$
Powell–Beale conjugate gradient	TRAINCGB	20.700	12	0.896	$y = 0.73x + 1.9$
Fletcher–Reeves conjugate	TRAINCGF	29.000	18	0.855	$y = 0.76x + 1.8$
Polak–Ribière conjugate gradient	TRAINCGP	28.600	10	0.818	$y = 0.59x + 2.7$
Gradient descent	TRAINGD	1500.000	6	0.405	$y = 0.47x + 40$
Gradient descent with momentum	TRAINGDM	706.000	6	0.130	$y = 0.22x + 20$
Gradient descent with adaptive learning rate	TRAINGDA	22.000	260	0.918	$y = 0.82x + 2.3$
Gradient descent with momentum and adaptive learning rate	TRAINGDX	105.000	29	0.327	$y = 0.1x + 6.6$
Levenberg–Marquardt	TRAINLM	0.024	6	1.000	$y = 0.99x + 0.013$
One-step secant	TRAINOSS	15.700	22	0.911	$y = 0.8x + 2.7$
Resilient (Rprop)	TRAINRP	44.800	21	0.821	$y = 0.67x + 1.2$
Scaled conjugate gradient	TRAINSCG	39.000	26	0.813	$y = 0.68x + 2.5$

3.2. Optimization of ANN Structure

The optimization of the ANN structure was determined based on the MSE value and the regression of the model. In the optimization, two neurons were used as the initial guess in the hidden layer for this network. When the number of neurons was increased, the MSE produced by the network became smaller. At a certain number of neurons, the MSE of the network started to increase from the minimum value of MSE. Figure 1 shows that the MSE of the network was the highest at 2 neurons with a value of 5.98. When the number of neurons was increased from 3 to 6, MSE decreased significantly from 5.35 to 0.188. The network achieved the minimum value of MSE of 0.0277 with 7 neurons at the hidden layer. When the number of neurons was increased from 7 to 14, the MSE of the network gradually increased from 0.0277 to 0.879. A further increase in the number of neurons from 14 to 15 resulted in a sharp increase in the MSE.

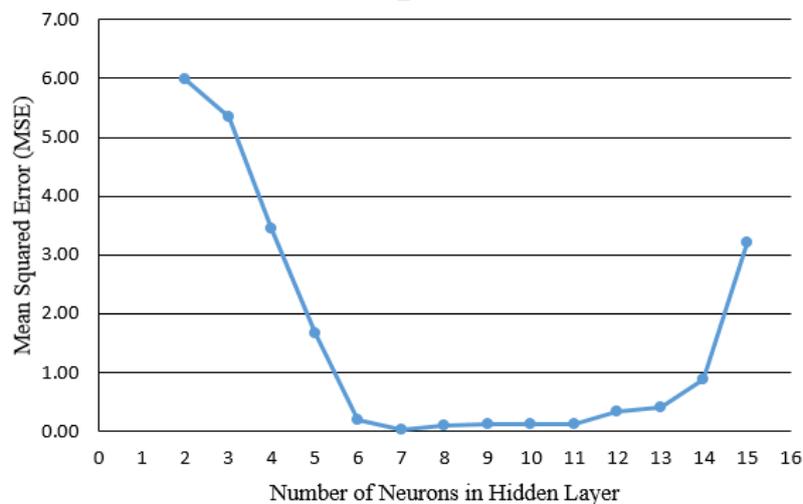


Figure 1. Relationship between MSE and number of neurons in the hidden layer.

At the same time, when the number of neurons is increased, the regression of the predicted outputs will increase. Figure 2 illustrates the dependence of regression and the number of neurons in the hidden layers. When the number of neurons was increased from 2 to 6, the regression of the predicted outputs increased from 0.9706 to 0.9992, but there is an exception when there are 3 neurons at the hidden layer where the regression decreased from 0.9706 to 0.9597 due to the combinatorial nature of the experimental data. The combinatorial experimental data sent to training was scattered which reduced the regression of the predicted output. The highest regression of predicted output of 0.9999 was achieved with 7 neurons at the hidden layer. A further increase in the number of neurons from 7 to 15 reduced the regression of the predicted output. The performance of the network is better when the number of neurons is increased. However, overfitting will occur when there is too much number of neurons. Overfitting will lead to the predicted outputs to be exact with or close to the experimental data and therefore may fail to fit additional data [17].

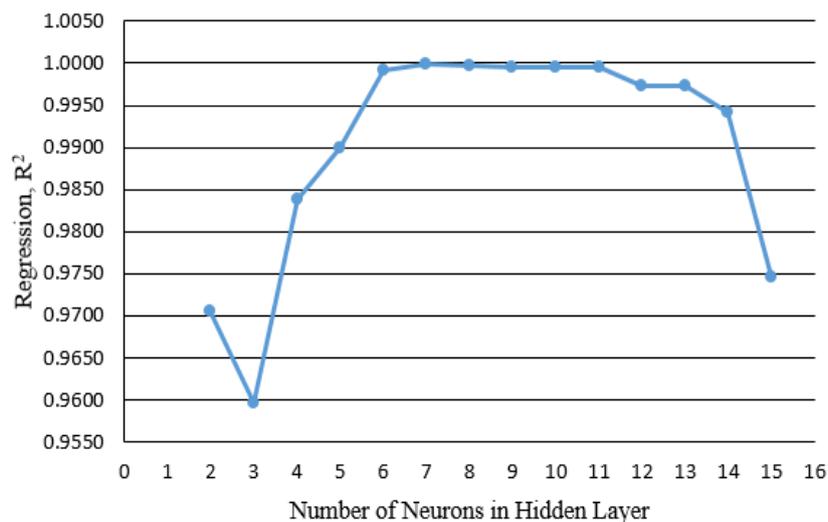


Figure 2. Relationship between regression and number of neurons in hidden layer.

The number of iteration determines the training of the network. When the predicted output deviates from the experimental data, the weight of the deviation will be adjusted, and the input of the network will be reinitialized and the network is trained again with the new input. Training will stop when the predicted output is close to the experimental data (the error was set to 1%). In this modeling, the training was stopped after 43 iterations out of 1000 (Epochs 43/1000) for LMA because further iterations show that the MSE started to increase. Figure 3 shows the MSE of training, validation, and test at different epochs for TRAINLM. The validation set was used to determine when to stop the iteration of the training set to prevent overfitting. The best epoch was at 37 because the MSE of validation started to increase after 37 epochs. At 37 epochs, the optimum MSE of the training was 0.0277.

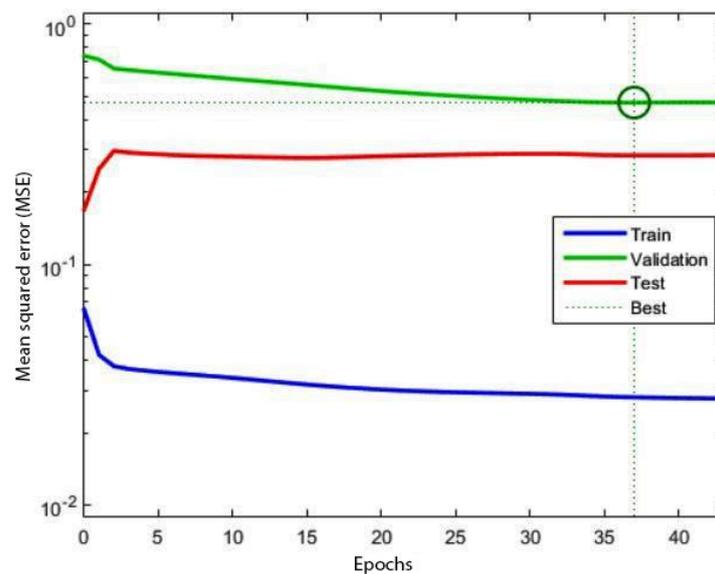


Figure 3. MSE of training, validation, and test against epochs.

Finally, the optimum artificial neural network with the BP algorithm was designed for this modeling. The network consists of three layers, which are input layer, hidden layer, and output layer. There are seven neurons at the hidden layer with tangent sigmoid transfer function (TANSIG) and an output layer

with linear transfer function (PURELIN). Figure 4 shows the optimum structure of ANN, together with the BP algorithm flow chart for the prediction of residual vanillin concentration.

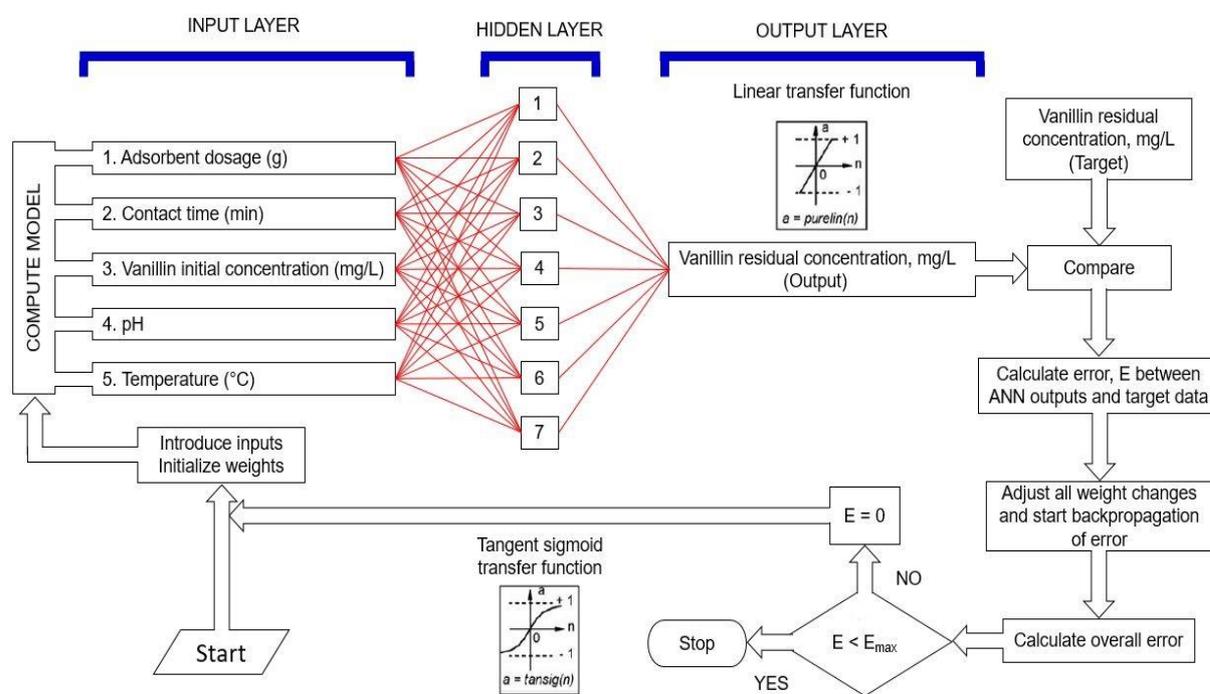


Figure 4. Optimum ANN structure with BP algorithm flow chart for prediction of residual vanillin concentration

3.3. Modeling of Vanillin Adsorption

In this modeling, the optimum ANN structure was used to test the performance of five variables which are adsorbent dosage, contact time, initial vanillin concentration, pH, and temperature. Their performance was evaluated based on the MSE of the training and regression of the model. A good model is a model with low MSE and high regression. For this modeling, the training network provided an MSE of 0.024 and the correlation of determination for the model is 0.9999, which proved that ANN is suitable to model the vanillin adsorption using resin H103. ANN is excellent in solving nonlinear model because it can produce a precise and effective response in a short time. The precision of the model was achieved by the BP algorithm. It calculated the error of the predicted output from the experimental data and propagated the error backwards, so that the input could be reinitialized and to be trained again. This process was repeated until the predicted output was close to the experimental data. On the other hands, LM algorithm which was used as a training transfer function can solve the nonlinear problem of the experimental data. It can minimize the problems of least-square curve fitting. Figure 5 shows the predicted residual vanillin concentration against the experimental residual vanillin concentration. The predicted vanillin concentration is very close and almost equal to the experimental residual concentration which means that the modeling of vanillin adsorption by ANN is successful [19].

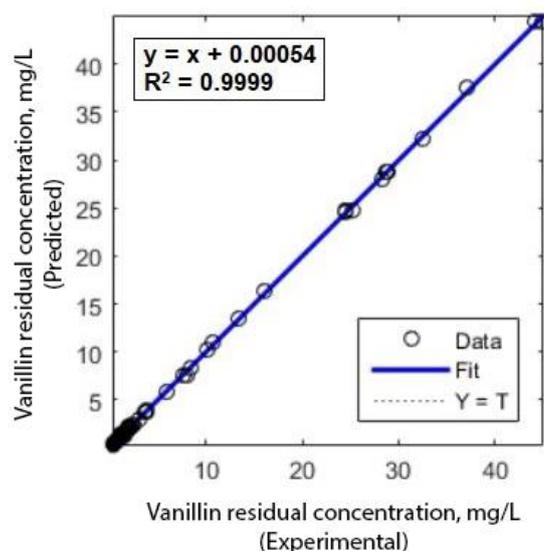


Figure 5. Predicted vanillin residual concentration versus experimental residual concentration.

3.4. Effect of Adsorbent Dosage on Adsorption Efficiency

The effect of adsorbent dosage on adsorption efficiency was studied by increasing the adsorbent dosage from 0.5 to 5.0 g while the contact time, initial vanillin concentration, pH, and temperature were maintained at 90 min, 50 mg/L, 5.94, and 25 °C, respectively. Figure 6 shows the efficiency between the experimental data and predicted output as a function of adsorbent dosage. The experimental data show that the adsorption efficiency increased from 92.37% to 98.11% when the adsorbent dosage was increased from 0.5 to 2.0 g. However, a further increase of adsorbent dosage to 5.0 g reduced the adsorption efficiency from 98.11% to 96.69%. According to Garg et al. [20], this can be attributed to the increase of surface area and the availability of adsorption sites for the removal of vanillin when the weight of the adsorbent is increased. However, a further increase of adsorbent dosage may cause aggregation and leads to overlapping of adsorption sites. Consequently, the available adsorption sites may decrease as well due to the adsorption density. The maximum adsorption efficiency was found to be 98.11% at 2.0 g of adsorbent dosage.

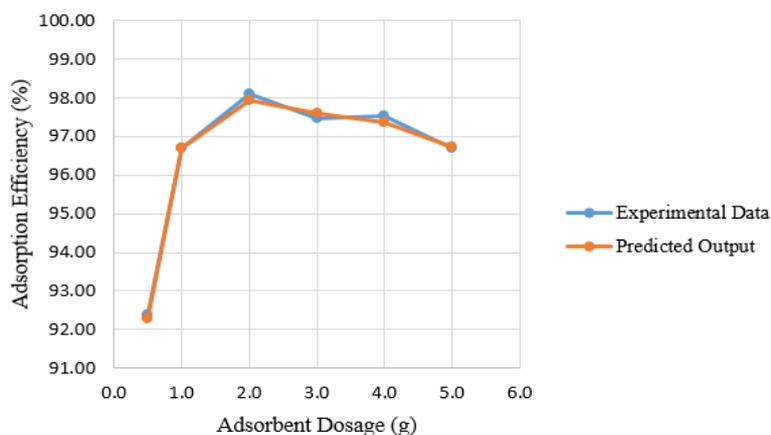
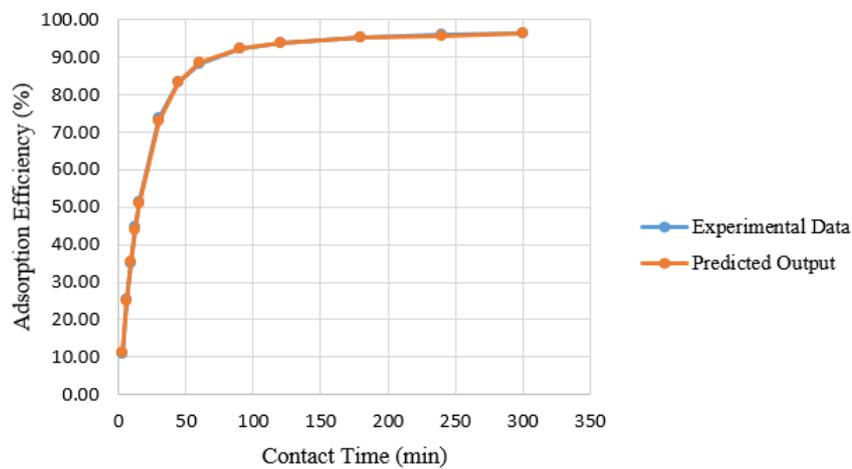


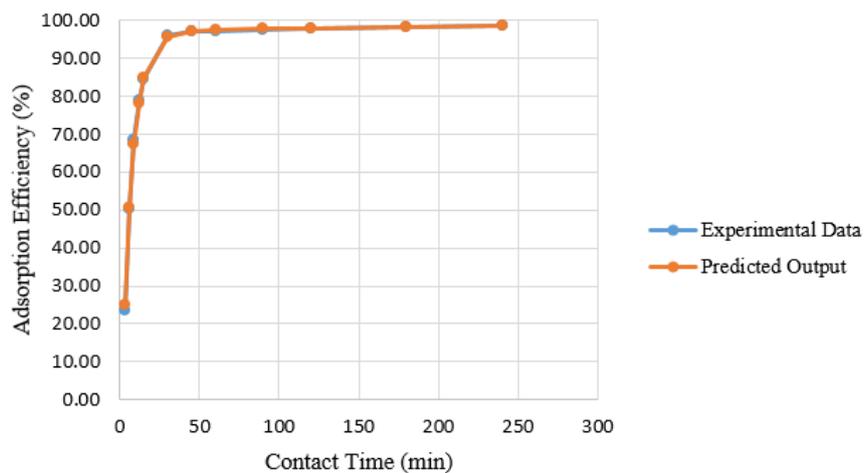
Figure 6. Adsorption efficiency between experimental data and predicted output as a function of adsorbent dosage

3.5. Effect of Contact Time on Adsorption Efficiency

The effect of contact time on adsorption efficiency was studied by increasing the contact time up to 300 min while the adsorbent dosage, initial vanillin concentration, pH, and temperature were maintained at 0.5 g, 50 mg/L, 5.94, and 25 °C, respectively. The experiment was repeated by replacing the adsorbent dosage to 2.0 g to compare the effect of adsorbent dosage and contact time on the adsorption efficiency. The experimental data show that a contact time of 60 and 30 min at 0.5 and 2.0 g of adsorbent dosage, respectively, are sufficient for the adsorption efficiency to achieve equilibrium (figure 7). A further increase of contact time did not yield a significant increase in the adsorption efficiency. Therefore, a further increase in contact time is unnecessary by considering the cost of energy consumed in agitation. The adsorption efficiency at 60 min with 0.5 g of adsorbent dosage was 88.17% while the adsorption efficiency at 30 min with 2.0 g of adsorbent dosage was 96.03%. The adsorbent dosage of 2.0 g was chosen because it could achieve higher adsorption efficiency with lesser energy consumed in agitation due to decrease in contact time.



(a)



(b)

Figure 7. Adsorption efficiency between experimental data and predicted output as a function of contact time for (a) 0.5 and (b) 2.0 g of adsorbent dosage

3.6. Effect of Initial Vanillin Concentration on Adsorption Efficiency

The effect of initial vanillin concentration on adsorption efficiency was studied by increasing the initial vanillin concentration from 6.25 to 800 mg/L while the adsorbent dosage, contact time, pH, and temperature were maintained at 0.5 g, 90 min, 5.94, and 25 °C, respectively. The experimental data in figure 8 show that the adsorption efficiency increased gradually from 93.14% to 98.14% when the initial vanillin concentration was increased from 6.25 to 100 mg/L. The adsorption efficiency decreased from 98.14% to 96.40% as a result of further increase of initial vanillin concentration because beyond 100 mg/L of initial vanillin concentration, the adsorbent was already saturated where all of the adsorption sites were fully occupied. Besides, Michael and Ayebaemi [21] reported that the adsorbate takes a shorter time to adsorb themselves to the adsorption sites at low concentration of adsorbate. In contrast, it takes a longer time when the concentration of adsorbate is increased because the adsorbate needs to diffuse to the adsorbent site through the intraparticle diffusion mechanism. The optimum initial vanillin concentration was found at 100 mg/L with 98.14% adsorption efficiency.

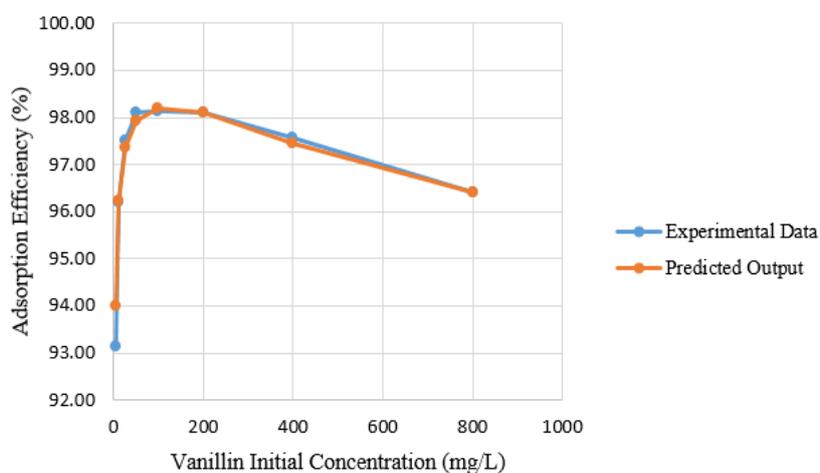


Figure 8. Adsorption efficiency between experimental data and predicted output as a function of initial vanillin concentration

3.7. Effect of pH on Adsorption Efficiency

The effect of pH on adsorption efficiency was studied by increasing the pH from 3.00 to 7.00 while the adsorbent dosage, contact time, initial vanillin concentration, and temperature were maintained at 0.5 g, 90 min, 50 mg/L, and 25 °C, respectively. The experimental data show that when pH was increased from 3.00 to 5.00, the adsorption efficiency slightly increased from 97.93% to 98.17% because at low pH, there is a high concentration of H^+ ions to compete with vanillin to bind themselves with the active sites. Therefore, when pH was increased, H^+ ions decreased, resulting in the more available active site for vanillin, thus increasing the efficiency of vanillin removal. However, no substantial difference in the adsorption efficiency within the tested pH range, which might be due to the neutral state of vanillin in the pH range of 2.0–6.0 [22]. The optimum pH was found to be 4.00 for this modeling. Figure 9 shows the comparison of efficiency between the experimental data and predicted output as a function of pH. Different from the other factors, the ANN model shows some deviation from the experimental data which might be caused by the insufficient experimental data to be trained in the ANN model which made the ANN model to fail to predict the experimental data accurately [23].

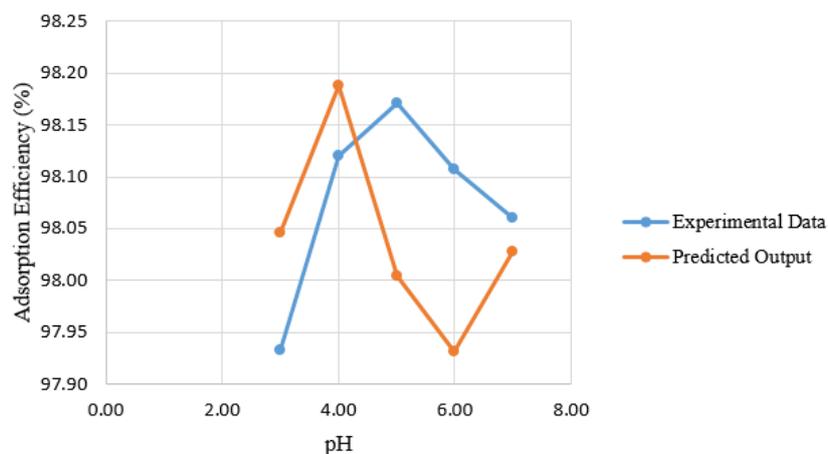


Figure 9. Adsorption efficiency between experimental data and predicted output as a function of pH.

3.8. Effect of Temperature on Adsorption Efficiency

The effect of temperature on adsorption efficiency was studied by increasing the temperature from 25 to 55 °C while the adsorbent dosage, contact time, initial vanillin concentration, and pH were maintained at 0.5, 90 min, 50 mg/L, and 5.94, respectively. The experimental data (figure 10) show that the adsorption efficiency kept decreasing from 98.10% to 96.03% when the temperature was increased from 25 to 55 °C. Because adsorption is a physisorption process [11], when temperature increases, vanillin molecules gain enough kinetic energy to overcome the electrostatic force of attraction between the adsorbent and themselves and escapes from the surface of the adsorbent which leads to a decrease in the adsorption efficiency. The optimum temperature for the adsorption process for this model was found to be 25 °C.

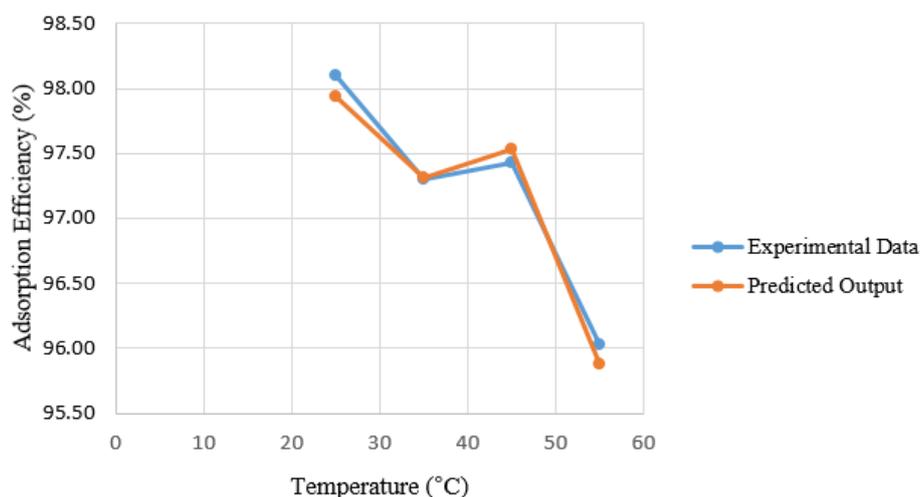


Figure 10. Adsorption efficiency between experimental data and predicted output as a function of temperature.

4. Conclusions

The optimum backpropagation ANN for this modeling is three layers with tangent sigmoid transfer function (TANSIG) at the hidden layer and linear transfer function (PURELIN) at the output layer. A minimum error of predicted data from experimental data was obtained by using the Levenberg–Marquardt algorithm with the MSE of only 0.024. The optimum ANN structure is with 7 neurons at the

hidden layer with a coefficient of determination of 0.9999 which means that the regression of the model is very high and the MSE value of 0.0277. The best adsorption efficiencies for each factor were 98.11%, 96.03%, 98.14%, 98.17%, and 98.10% at 2.0 g of adsorbent dosage, 30 min (2.0 g of adsorbent dosage), 100 mg/L of initial vanillin concentration, pH 5, and 25 °C, respectively. From the result, ANN shows an excellent performance in predicting the experimental data and suitable to be used in the modeling of vanillin adsorption.

Acknowledgment

We would like to thank you Universiti Malaysia Pahang for providing MATLAB for the modeling throughout the research.

References

- [1] Zhang Q F, Jiang Z T, Gao H J and Li R 2008 *Eur. Food Res. Technol.* **226**(3) 377–83
- [2] Mota M I F, Pinto P C R, Loureiro J M and Rodrigues A E 2016 *Sep. Purif. Rev.* **45**(3) 227–59
- [3] Chambers C M 1997 *Proc. of The Internal Structure of Black Holes and Spacetime Singularities* 490–96
- [4] Lomascolo A, Stentelaire C, Asther M and Lesage-Meessen L 1999 *Trends Biotechnol.* **17**(7) 282–89
- [5] Box G E P and Behnken D W 1960 *Technometrics* **2**(4) 455–75
- [6] Ba D and Boyaci I H 2007 *J. Food Eng.* **78**(3) 836–45
- [7] Elemen S, Akçakoca Kumbasar E P and Yapar S 2012 *Dye. Pigment.* **95**(1) 102–11
- [8] Kasiri M B, Aleboyeh H and Aleboyeh A *Environ. Sci. Technol.* **42**(21) 7970–75
- [9] Singh V, Khan M, Khan S and Tripathi C K M 2009 *Appl. Microbiol. Biotechnol.* **82**(2) 379–85
- [10] Gevrey M, Dimopoulos I and Lek S 2003 *Ecol. Modell.* **160**(3) 249–64
- [11] Samah R A 2016 *Optimization of Batch Adsorption and Fixed-Bed Adsorption of Vanillin onto Resin H103*. Universiti Putra Malaysia, Selangor, Malaysia.
- [12] Ortega-Zamorano F, Jerez J M, Munoz D U, Luque-Baena R M and Franco L 2016 *IEEE Trans. Neural Networks Learn. Syst.* **27**(9) 1840–50
- [13] Karnin E D 1990 *IEEE Trans. Neural Networks* **1**(2) 239–42
- [14] Gueguim Kana E B, Oloke J K, Lateef A and Adesiyun M O 2012 *Renew. Energy* **46** 276–81
- [15] Basheer I A and Hajmeer M 2000 *J. Microbiol. Methods* **43**(1) 3–31
- [16] Mele B and Altarelli G 1993 *Phys. Lett. B* **299**(3–4) 345–50
- [17] Kumar K V and Porkodi K 2009 *Chem. Eng. J.* **148**(1) 20–25
- [18] Yetilmezsoy K and Demirel S 2008 *J. Hazard. Mater.* **153**(3) 1288–300
- [19] Satoh H and Namerikawa T 2007 *Proc. IEEE Int. Conf. Control Appl.* 539–44
- [20] Garg V K, Gupta R, Yadav A B and Kumar R 2003 *Bioresour. Technol.* **89**(2) 121–24
- [21] Michael H J and Ayebeami I S 2005 *Electron. J. Biotechnol.* **8**(2) 162–69
- [22] Li R, Jiang Z, Mao L and Shen H 1998 *Talanta* **47**(5) 1121–27
- [23] Demiral Y F and Öztürk N 2013 *Proc. of ICOEST* 246–55