Quality Prediction Modeling of Palm Oil Refining Plant in Malaysia Using Artificial Neural Network Models

Nurul Sulaiha Sulaiman[1], Khairiyah Mohd-Yusof[2*], Asngari Mohd-Saion[3]

[1]*School of Chemical and Energy Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, Malaysia*
[2]*Center for Engineering Education, Universiti Teknologi Malaysia, Malaysia*
[3]*Sime Darby Kempas Sdn Bhd, Pasir Gudang, Johor Bahru, Malaysia*
Corresponding author E-mail:khairiyah@utm.my

Abstract

Malaysia is currently one of the biggest producers and exporters of palm oil and palm oil products. In the growth of palm oil industry in Malaysia, quality of the refined oil is a major concern where off-specification products will be rejected thus causing a great loss in profit. In this paper, predictive modeling of refined palm oil quality in one palm oil refining plant in Malaysia is proposed for online quality monitoring purposes. The color of the crude oil, Free Fatty acid (FFA) content, bleaching earth dosage, citric acid dosage, activated carbon dosage, deodorizer pressure and deodorizer temperature were studied in this paper. The industrial palm oil refinery data were used as input and output to the Artificial Neural Network (ANN) model. Various trials were examined for training all three ANN models using number of nodes in the hidden layer varying from 10 to 25. All three models were trained and tested reasonably well to predict FFA content, red and yellow color quality of the refined palm oil efficiently with small error. Therefore, the models can be further implemented in palm oil refinery plant as online prediction system.

Keywords: *Artificial Neural Network; Edible Oil Refining; Palm Oil; Quality Prediction*

1. Introduction

Crude Palm Oil (CPO) needs to undergo refining process to remove impurities and meet specific quality requirement especially in FFA content and its color. Bleaching, degumming and deodorization are three major steps in this refining process. To avoid color fixation in deodorization step, phosphorus and iron are needed to be reduced to sufficiently low level in degumming step [1] while bleaching step removed the color pigments in the CPO by physical adsorption on the bleaching agent surface [2]. The final product is called Refined, Bleached and Deodorized Palm Oil (RBDPO).

In the overall refining process, several variables affect the process performance and the RBDPO quality. The most crucial variables that affect the RBDPO quality are listed as quality of the CPO [3], acid concentration [4], amount of adsorbent used [5, 6, 7], temperature and vacuum pressure of the process [8]. In the current palm oil refining industry, the quality of the RBDPO is manually monitored by chemical analysis in laboratory which is very tedious, expensive and time consuming. If the RBDPO does not meet the specific quality requirements, it will be labelled as off-spec product and need to be refined again. This will cause increase in the cost as well as time delay in the production.

On the other hand, with this large number of variables affecting the refining process, mathematical model development of the process is an exceptionally complex task. The use of ANN algorithms is an alternative to overcome the mathematical modeling problem. ANN possess the ability to learn what happens in the process without actually modeling the physical and chemical laws that govern the system which also called black box approach [9]. Therefore, ANN has been widely used in the application of chemical engineering due to its ability to act as a black-box systems for data where the underlying relation are poorly known or extremely complete [10].

The uses of ANN in palm oil related work include control in palm oil mills [11], prediction of palm oil properties [12], forecasting oil palm yield [13], and latest work used ANN in predicting RBDPO quality of FFA content [14, 15], concentration of beta-carotene and tocopherol [15] and prediction of phosphoric acid and bleaching earth consumption [16, 17].

The main objective of this study is to develop ANN models that can be used in online monitoring of RBDPO qualities. The developed ANN models in the study are able to predict the final qualities of the RBDPO which are FFA content and color of the RBDPO. For this purpose, ANN models were trained and validated using FFA content and color of the CPO, bleaching earth dosage, citric acid dosage, activated carbon dosage, deodorizer pressure and deodorizer temperature as the input to the model. A feed forward neural network was designed using a Levenberg-Marquardt (LM) training algorithm. The ability of ANN to model the palm oil refining process is demonstrated in this paper.

The paper is organized in the following way; in Section 2, industrial palm oil refining process in one of palm oil refinery plant in Malaysia is described. Subsequently, in Section 3, the predictive modeling of palm oil refining process using ANN is presented. Finally, the results and comparison of the ANN models are presented in Section 4 and the conclusion reached in this study is discussed in Section 5.

2. Description of Industrial Palm Oil Refining Process

Data used in this study is obtained from one palm oil refinery plant in Malaysia. When crude oil entered the bleacher unit, the citric acid is dosed and degumming process takes place. Citric acid coagulates impurities like phosphatides, chlorophyll, carotenoid, trace metal ions and oxidative product. Next, bleaching earth and activated carbon as bleaching agent are dosed to the oil. The bleaching process retention time is about 30 minutes. Bleaching earth absorbed the coagulated impurities and later it is removed by filtration process. The bleached palm oil then undergone deodorization process where all volatile components such as FFA and odoriferous pigments are distilled and removed as Palm Fatty Acid Distillate (PFAD). This process is also known as distillation process (temperature range: 245-265°C, vacuum pressure).

In current practice, the final quality of the RBDPO is manually monitored through lab testing. This method is very tedious and time consuming. The RBDPO that does not meet the quality requirement will be labelled as off-spec product and need to be refined again. This cause increase in production cost as well as delay in the production. Thus, this research is focusing on predicting the quality of RBDPO.

3. Model Development

3.1. ANN modelling Strategy

In this study, a feed forward neural network was utilized to predict the final quality of the RBDPO. The models developed consist of an input layer, one hidden layer and an output layer. Five single ANN models were developed based on the output selected. MISO-1, MISO-2 and MISO-3 are single prediction model for single output which are FFA content, red Lovibond color and yellow Lovibond color respectively. MIMO-1 is a prediction model for multiple output of red and yellow Lovibond color while MIMO-2 is a prediction model for multiple output of FFA, red and yellow Lovibond color. The ANN models are listed in Table 1.

Table 1: Network group

ANN Model	Prediction
MISO-1	FFA content
MISO-2	Red Lovibond color
MISO-3	Yellow Lovibond color
MIMO-1	Red and Yellow Lovibond color
MIMO-2	FFA content, red and yellow Lovibond color

3.2. Industrial Data Acquisition

A total of 100 data set referring to the final quality of the RBDPO were collected. The process variables considered in this study were: FFA content, color, bleaching earth, activated carbon and citric acid dosage, as well as temperature and pressure of the deodorizer. These variables are set as input data to the models while the output data is the final quality of RBDPO which are FFA content and color. All input and output variables are listed in Table 2. Before the CPO entered bleaching unit, 100ml of the oil was taken and the quality of it was checked by chemical analysis in the laboratory and recorded. During degumming, bleaching and deodorization process, process conditions such as bleaching earth dosage, citric acid dosage, activated carbon dosage, temperature and pressure of the unit were recorded. After the oil came out from the deodorizer unit, the quality of the oil is again checked by chemical analysis technique and recorded.

Table 2: List of input and output variables

Number	Input variables
1	FFA content of CPO (%)
2	Red Lovibond color of CPO
3	Yellow Lovibond color of CPO
4	Blue Lovibond color of CPO
5	Bleaching earth dosage (%)
6	Citric acid dosage (%)
7	Activated carbon dosage (%)
8	Vacuum pressure of deodorizer (bar)
9	Sparging tray 1 temperature (°C)
10	Sparging tray 2 temperature (°C)
Number	Output variables
1	FFA content of RBDPO (%)
2	Red Lovibond color of RBDPO
3	Yellow Lovibond color of RBDPO

3.3. Pre-Processing Data

Normalizing or standardizing the inputs can make training faster and reduce the chances of getting stuck in local optima. The input and the output data are normalized within a uniform range of [−1 1] by using mapminmax function in Matlab.

3.4. Training and Testing Model

Before being introduced to the network, the data were divided into three parts; training, testing and validation. To optimize the number of hidden nodes, the trial-and-error method is used. The number of hidden node is changed in a manner from 10 to 15, then change to 20 and lastly 25 in order to improve the fitting of the process data. The transfer functions for both hidden and output layers were also determined as log–sigmoid and linear function, respectively. The output layer consists of node related to the final quality of the refined palm oil. The network is trained by using back-propagation (BP) algorithms and the weights and biases are iteratively updated using LM algorithm until the convergence to the certain value is achieved. To test the effectiveness the model, the performance of the network is calculated based on the error function, Mean Square Error (MSE) and regression (R).

4. Result and Discussion

In this study, the ANN is used to estimate the final quality of FFA content and color of the RBDPO. As mentioned in the previous section, there are five ANN models developed in this study. The models labeled as MISO-1, MISO-2 and MISO-3 are for single output prediction of FFA, red color and yellow color respectively. Another two models are which labelled as MIMO-1 and MIMO-2 are for colors (red and yellow) prediction and combined FFA and colors of RBDPO prediction respectively. The performance of all models were assessed by Regression (R) and Mean Squared Error (MSE) of training, testing, validation and overall set. The number of hidden nodes are changed simultaneously from 10, 15, 20 and lastly 25. R and MSE of each network were recorded in Table 3.

As shown in Table 3, varying the number of hidden nodes will results in different accuracy of the prediction. The highlighted results are the best R and MSE for all tested network. It is clearly shown that for MISO-1 which is single output network for FFA prediction, small number of hidden nodes which is 10 is sufficient enough to give a good prediction accuracy. The best network for MISO-1 is network with 10 number of nodes that give the smallest values of R and MSE which are 0.5764 and 0.0552 respectively. For MISO-2 and MISO-3 which predict red color and yellow color of RBDPO respectively, with number of hidden nodes of 20, each network   gave

**Table 3:** Regression (R) and Mean Squared Error (MSE) for all network

Network	No. of Nodes	R				MSE			
		Training	Validation	Testing	All	Training	Validation	Testing	All
	10	**0.5503**	**0.8007**	**0.7200**	**0.5764**	**0.0630**	**0.0275**	**0.0200**	**0.0552**
MISO-1	15	0.4480	0.1669	0.1548	0.3647	0.1826	0.1794	0.5129	0.2153

|								
	20	0.4871	0.1212	0.5713	0.4421	0.1122	0.2640	0.0715	0.1233
	25	0.0313	0.0237	-0.3798	0.0173	0.1717	0.2013	0.1583	0.1733
MISO-2	10	0.8366	0.8250	**0.7826**	0.8122	0.0787	0.0985	**0.1908**	0.0919
	15	0.9362	0.6231	0.5033	0.8400	0.0309	0.1833	0.3551	0.0785
	20	**0.9407**	**0.8450**	0.6725	**0.8845**	**0.0283**	**0.1000**	0.2628	**0.0589**
	25	0.9345	0.5691	0.5642	0.8012	0.0313	0.4004	0.3771	0.1028
MISO-3	10	0.5178	0.8236	**0.8214**	0.5597	0.2709	0.1543	**0.2776**	0.2599
	15	**0.9539**	0.6632	0.3756	0.8301	**0.0253**	0.2546	0.5124	0.0970
	20	0.8945	**0.8373**	0.6373	**0.8480**	0.0658	**0.0811**	0.2964	**0.0904**
	25	0.9411	0.5589	0.4977	0.7927	0.0321	0.5154	0.4720	0.1244
MIMO-1	10	0.4455	0.3286	0.1105	0.3751	0.2805	0.2364	0.5136	0.2994
	15	0.8638	**0.8249**	0.6006	0.8093	0.0771	**0.2123**	0.3307	0.1160
	20	0.8467	0.6503	0.6392	0.7650	0.0839	0.2228	**0.2874**	0.1181
	25	**0.9213**	0.6705	**0.7176**	**0.8554**	**0.0556**	0.3034	0.3057	**0.1054**
MIMO-2	10	**0.9578**	0.7082	0.6793	**0.8926**	**0.0213**	0.1471	**0.2253**	**0.0543**
	15	0.9182	**0.8141**	0.5658	0.8537	0.0415	**0.1139**	0.3100	0.0756
	20	0.6354	0.4960	0.5361	0.5898	0.1869	0.2667	0.2979	0.2060
	25	0.9517	0.5967	**0.6843**	0.8702	0.0250	0.2715	0.2397	0.0711

an excellent prediction with 0.8845 and 0.8480 values of R and 0.0589 and 0.0904 values of MSE respectively. This shows that both network have similar mapping of input to the predicted output.

For MIMO-1 and MIMO-2 network, there is a huge difference for both network s. MIMO-1 which is for prediction of both red and yellow color of the RBDPO with number of hidden nodes of 25 gave the best result for both R and MSE values which are 0.8554 and 0.1054 respectively. On the other side, network MIMO-2 for prediction of all FFA, red and yellow color with the smallest number of hidden nodes (10) gave the best result of R and MSE values. This clearly proved that the output of FFA gave a greater effect toward the mapping of input and output of the network.

Comparing all R and MSE for each training, validation, testing and overall set, training set gave a superb result of both R and MSE for all networks except for MISO-1. This is due to the data division; 80% of the total data are training data, while only 20% of the total data which 10 % each is for validation and testing set respectively. This concludes that, a larger data point gave an outstanding prediction for each network. The regression of each network for prediction performance is shown in Fig. 1(a)-(e).
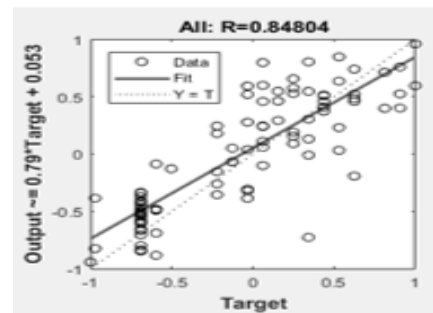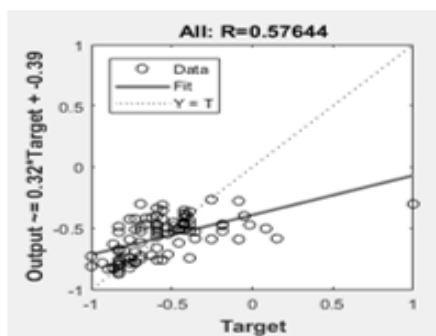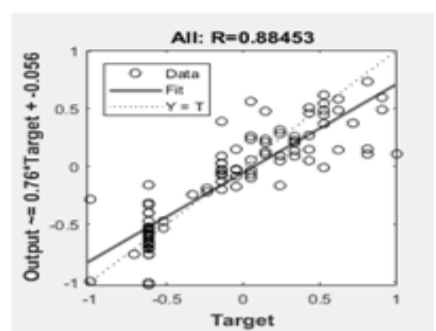

(c)   MISO-3
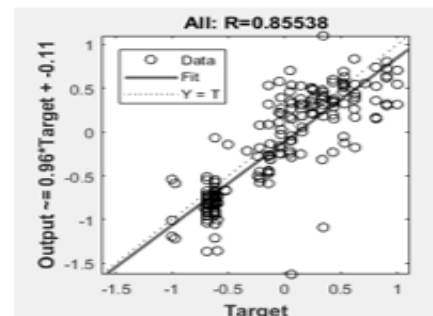
(d)   MIMO-1
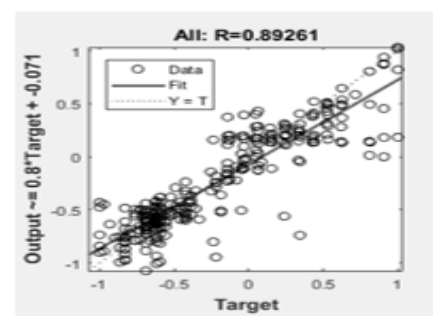
(a)   MISO-1

(b)   MISO-2

(e)   MIMO-2
Fig.1: Regression of all network

Fig. 1 (a) – (e) shows the regression performance of all predictions by each ANN models which has a correlation of more than 80 % with the actual process parameters for all model except MISO-1. The distribution of data hugely affect the prediction performance. As can be seen in the figure, the best network is MIMO-2 which predicts all FFA, red and yellow color of the RBDPO. The large data point, gave a good training and thus gave an outstanding prediction compared to single network MISO-1, a small variation in data resulted in low quality of training and thus poor prediction of FFA. Comparing single output and multiple output network s, multiple output networks (MIMO-1 and MIMO-2) gave a superior prediction of output compared to all single output networks (MISO-1, MISO-2 and MISO-3). To increase the prediction perfor-

mance of MIMO-1 and MIMO-2, Stack Neural Network (SNN) are developed. The R and MSE values of each network is shown in Table 4.

Table 4: Regression (R) and Mean Squared Error (MSE) for SNN network

No. of nodes	R		MSE	
	SNN-MIMO-1	SNN-MIMO-2	SNN-MIMO-1	SNN-MIMO-2
10	0.9998	0.9996	8.7162e-05	2.1544e-04
15	0.9999	0.9997	7.7442e-05	1.5200e-04
20	0.9999	0.9998	7.6382e-05	8.8721e-05
25	0.9999	0.9999	3.9546e-05	7.0348e-05

SNN network for both MIMO-1 and MIMO-2 gave a better result compared to its ANN network. The regression of almost 1.0 recorded for both network showed an absolutely perfect prediction of output for each network. Network with the smallest value of MSE is network with 25 number of hidden nodes for both SNN-MIMO-1 and SNN-MIMO-2 which are 3.9546e-05 and 7.0348e-05 respectively. It can be concluded that the best network is network with 25 hidden nodes.

5. Conclusion

From this study, the prediction network was developed based on actual palm oil refining plant data using ANN model. The result showed that each network with different number of hidden nodes gave different prediction performance of both R and MSE. Different regression of each network is due to different data distribution of each network and larger data point division of training gave a good training of the network thus gave an outstanding prediction of the output. Multiple output network (MIMO-1 and MIMO-2) gave a better prediction accuracy compared to all single output networks (MISO-1, MISO-2 and MISO-3). Finally, SNN network gave a perfect prediction for both MIMO-1 and MIMO-2.

Acknowledgement

References

[1] Gibon, V., Greyt, D. W., and Kellens, M., "Palm oil refining", *European Journal of Lipid Science and Technology,* Vol 109(4), (2007), pp. 315-335.

[2] Silva, S. M., Sampaio, K. A., Ceriani, R., Verhé, R., Stevens, C. Greyt, W. D. and Meirelles, A. J. A., "Adsorption of carotenes and phosphorus from palm oil onto acid activated bleaching earth: Equilibrium, kinetics and thermodynamics", *Journal of Food Engineering,* Vol.118, (2013), pp. 341–349.

[3] Basiron, Y., Palm Oil, in: Shahidi, F. (Ed.), Bailey's Industrial Oil and Fat Products, 6 ed. John Wiley & Sons, Inc., (2005), pp. 333 - 429.

[4] Hussin, F., Aroua, M. K. and Wan-Daud, W. M. A., "Textural characteristics, surface chemistry and activation of bleaching earth: A review." *Chemical Engineering Journal*, Vol.170, (2011), pp. 90–106.

[5] Puah, C. W., Choo, Y. M., Ma, A. N., and Chuah, C. H., "Degumming and Bleaching: Effect on Selected Constituents of Palm Oil", *Journal of Oil Palm Research*, Vol.16 (2), (2004), pp. 57-63.

[6] Skevin, D., Domijan, T., Kraljic, K., Kljusuric, J. G., Nederal, S. and Obranovic, M., "Optimization of Bleaching Parameters for Soybean Oil", *Food Technology Biotechnology*, Vol.50 (2), (2012), pp. 199-207.

[7] Taylor D. R., "Bleaching", John Wiley & Sons, Inc., (2005), pp.321-330

[8] Greyt W. D., and Kellens, M., "Deodorization", John Wiley & Sons, Inc., (2005), pp. 321-330.

[9] Almeida, J. S., "Predictive non-linear modeling of complex data by artificial neural networks", *Current Opinion in Biotechnology*, Vol.13, (2002), pp. 72–76.

[10] Saxen, H. and Pettersson, F., "Method for the Selection of Inputs and Structure of Feedforward in Neural Networks", *Computer and Chemical Engineering*, Vol.30, (2006), pp.1038-1047.

[11] Ahmad, A.L., Azid, I. A., Yusof, A.R., and Seetharamu, K.N., "Emission Control in Palm Oil Mills using Artificial Neural Network and Genetic Algorithm", *Computers & Chemical Engineering*, Vol.28(12), (2004), pp. 2709-2715.

[12] Abdullah, S. and Eow C. T., "Prediction of Palm Oil Properties using Artificial Neural Network", *International Journal of Computer Science and Network Security*, Vol.8(8), (2008), pp. 101-106.

[13] Ismail, Z. and Khamis, A., "Neural Network in Modeling Malaysian Oil Palm Yield", *American Journal of Applied Sciences* Vol.8 (8), (2011), pp. 796-803.

[14] Sulaiman N. S. and Mohd-Yusof, K., "Artificial neural network-based model for quality estimation of refined palm oil," *2015 15th International Conference on Control, Automation and Systems (ICCAS)*, Busan, (2015), pp. 1324-1328.

[15] N. Tehlah, P. Kaewpradit, I.M. Mujtaba, "Artificial neural network based modeling and optimization of refined palm oil process", *Neurocomputing*, Vol.216, (2016), pp. 489-501.

[16] Morad, N. A., Zin, R. M., Yusof, K. M., and Aziz, M. K. A., "Process Modelling of Combined Degumming and Bleaching in Palm Oil Refining Using Artificial Neural Network", *JAOCS, Journal of the American Oil Chemists' Society*, Vol.87(11), (2010), pp. 1381-1388.

[17] Abdul-Hadi N. and Che-Mat, C.R, "Prediction Model Based on Artificial Neural Network for Industrial Bleaching and Degumming Process of Crude Palm Oil" *International Journal of Engineering and Innovative Technology (IJEIT)*, Vol.6 (1), (2016), pp. 32-38.