CHEMICAL ENGINEERING TRANSACTIONS
VOL. 114, 2024
A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.cetjournal.it
Guest Editors: Petar S. Varbanov, Min Zeng, Yee Van Fan, Xuechao Wang

ISBN 979-12-81206-12-0; **ISSN** 2283-9216
DOI: 10.3303/CET24114089

Optimisation of Crystallisation Recipe for Varied Cloud Points Characteristics in Palm Oil Fractions

John Ting Zhi Zhang[a], Jeng Shiun Lim[b,*]

[a]Faculty of Chemical and Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[b]Process Systems Engineering Centre (PROSPECT), Research Institute for Sustainable Environment, Universiti Teknologi Malaysia (UTM), 81310 UTM Johor Bahru, Johor, Malaysia
jslim@utm.my

The management of product quality in palm oil crystallisation poses a formidable challenge. Although various model-based optimisation control strategies have been widely applied, their effectiveness hinges on understanding the intricate and highly nonlinear dynamic behavior of crystallisation. Notably, existing research has predominantly focused on diverse applications, such as wastewater treatment, sugar cane crystallisation, and the pharmaceutical industry, leaving a notable research gap in the crystallisation processes specific to the palm oil industry. This research attempts to fill this gap by investigating the impact of an optimisation tool that combines artificial neural network and genetic algorithm (ANN-GA) to optimize the crystallisation recipe, specifically the cooling segments of palm oil, for three different cloud points of palm olein (CP 6, CP 8, and CP 10). The artificial neural network (ANN), which uses the Levenberg-Marquardt algorithm, serves as an internal model for predicting process output, whereas the genetic algorithm (GA) investigates a wide range of recipe combinations to maximise yield. Using MATLAB for optimisation, the ANN-GA approach goes through training, testing, and validation steps with industry-derived datasets. The results show root mean sqaure error (RMSE) of 0.8411 for CP 6, 0.4317 for CP 8, and 0.4105 for CP 10, indicating that ANN is sensitive to dataset volumes. Using GA as an optimisation tool, it generates optimal input variables for industrial validation. Validation results reveal an enhanced yield of 63 % for CP 6 palm olein, 74 % for CP 8 palm olein, which is within industrial range (66-76 %), and 77.26 % for CP 10 palm olein, which is within the range of 76-79 %. Overall, the ANN-GA technique is effective in predicting complicated systems such as palm olein and palm stearin crystallisation processes.

1. Introduction

Palm oil is one of the most traded oils for food applications on the global oils and fats market. Palm oil is extracted from the outer pulp (mesocarp) of oil palm fruits (Elaeis guineensis). The mesocarp contains around 49 % palm oil, whereas the kernel contains approximately 50 % palm kernel oil (Omar et al., 2015). This has resulted in several studies on nutritional, food design and formulation, and crystallisation behaviour. In 2023, the world production of palm oil reached 79.22 million MT from 58.4 MT in 2013 (Brodnitz, 2022). The high demand makes it the world's main commodity crops. In Malaysia, the drastic increase in palm oil production becomes one of the main factors for 79.22 Mt of world production (Brodnitz, 2022).

In the 1960s and 1970s, advances in technology and transportation infrastructure further contributed to the growth of the palm oil industry. The introduction of high-yielding oil palm varieties and mechanized harvesting and processing techniques helped to increase productivity and reduce costs. One of the techniques that is widely used is crystallisation process. Crystallisation process is an important process in the production of biofuels, cosmetics, and cooking oil. It is often carried out on the industrial scale by batch cooling in agitated vessels and the need for some control over this process has long been recognised (Mao et al., 2023). The process aims to remove solid impurities such as wax and fatty acids to promote formation of different cloud points of palm olein that exhibit unique characteristics.

Paper Received: 18 June 2024; Revised: 14 August 2024; Accepted: 18 November 2024
Please cite this article as: Zhang J.T.Z., Lim J.S., 2024, Optimisation of Crystallisation Recipe for Varied Cloud Points Characteristics in Palm Oil Fractions, Chemical Engineering Transactions, 114, 529-534 DOI:10.3303/CET24114089

Crystallisation process can be done in several forms, depending on the specific application and desired outcomes (MacWilliams et al., 2024). Fractionation process usually follows after crystallisation of palm oil. Due to the highly competitive nature of the today industry, hence, efficient, and robust model is implemented to achieve excellent productivity in yield improvement and energy saving. There have been 2 main routes of modelling which are analytical modelling that is based on first principle rule and data driven modelling that is based on process data. Analytical modelling has been the traditional way of modelling since many years, but the data driven modelling has been taking over due to advancement in computational resources (Galvanauskas et al., 2006). One of the most successful data driven modelling is artificial neural network (ANN) to optimise process due to its excellent performance in generalisation, short computational running time and modelling of nonlinear systems without any prior knowledge (Cristea et al., 2003).
Moreover, ordinary optimisation techniques like sequential quadratic programming cannot locate the solution since the challenge of determining the optimal value is highly nonlinear and has numerous local optima (Drissi et al., 2024). Consequently, optimisation problem needs an optimisation method that can identify the global optimum using Genetic Algorithm (GA) as it is well known for its robustness to find optimal or near-optimal solutions. It is also because GA is able to mimic the evolutionary process of nature and follow the principle of survival of the fittest during the evolution (Drissi et al., 2024). The goal of this study are to identify the optimal crystallisation recipe that produce highest yield of palm olein during crystallisation process.

2. Methodology

In this study, different set of data were collected from the control room. The inputs were set as temperature of water and cooling duration at each segment, quality of refined, bleached, deodorized palm oil (RBDPO) whereas for the output was yield of palm olein produced. After collection, the combined artificial neural network and genetic algorithm (ANN-GA) was developed using R2022a MATLAB software. The first stage for the development of ANN was to determine the number of hidden layers that can yield minimum mean square error using loop of training, testing and validation stages. The loop ended when the hidden layer with the least MSE is achieved. The second stage would be the development of GA. GA would undergo mutation, crossover to select the best candidate. The fitness function was using the best network that had the lowest mean square error. Optimisation ended when the iteration reached maximum generations. The final result would be the optimal recipe or optimal input generated from GA. Figure 1 depicts the overall methodology of the study.

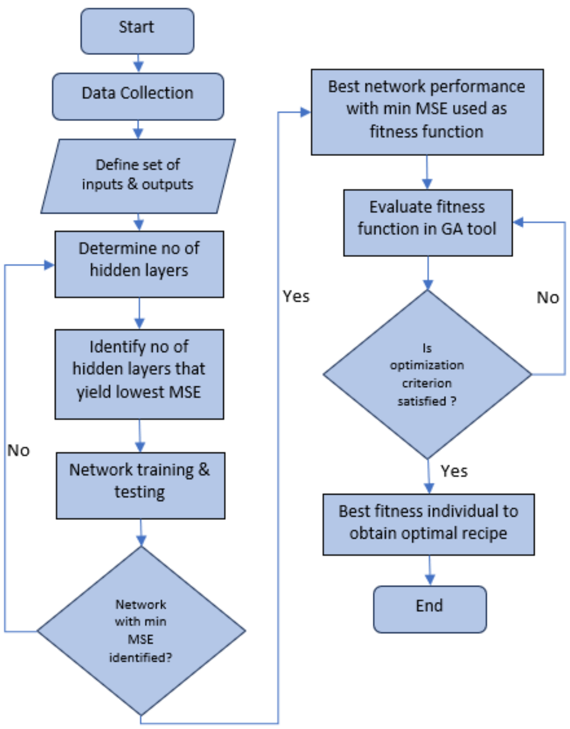

Figure 1: Flowchart for overall methodology

2.1 Background of the study

In the research, the crystallisation recipe, known as the cooling water segment is the manipulated variables used to control crystallisation process of palm olein and palm stearin. Water is the cooling medium while for RBDPO is the material that will be cooled and crystallised. There are 3 basic steps for crystallisation of fats which are supercooling of the melt, formation of crystal nucleic and crystal growth. Therefore, it is necessary to control the water temperature to resolve palm oil into various grades of palm olein and palm stearin.
Cooling curve is another term to describe the recipe used for crystallisation process. There are 4 stages of crystallisation which are divided into initial cooling, critical cooling, crystallisation period and final cooling. Before initial cooling, the temperature of palm oil is heated up to 70°C to destroy the crystal memory (Timms et al., 2005). When the temperature is reached, initial cooling will start to prepare the oil in supersaturated condition. In this condition, the oil temperature will start to drop slightly. Next, small crystals start to appear when the critical point is reached. The slope of critical cooling is steeper than initial cooling. Each cloud point of palm olein produced will have a similar critical point which is around 47-49°C. Crystallisation period is the time for crystals to agglomerate and form bigger crystals (Timms et al., 2005). During this period, there may be a small bump which indicates an exothermic reaction has occurred. This may be one of the factors that cause the crystal to have rough appearance and cause harder separation of palm olein and palm stearin (De Greyt et al., 1996). The last stage is final cooling in which the crystal will be further cooled to form a round and spherical appearance to prevent the fraction of olein being trapped between crystals during filtration. Figure 2 depicts the crystallisation curve that shows initial cooling, critical cooling, crystallisation period and final cooling phases in numberings 1,2, 3 and 4 respectively.

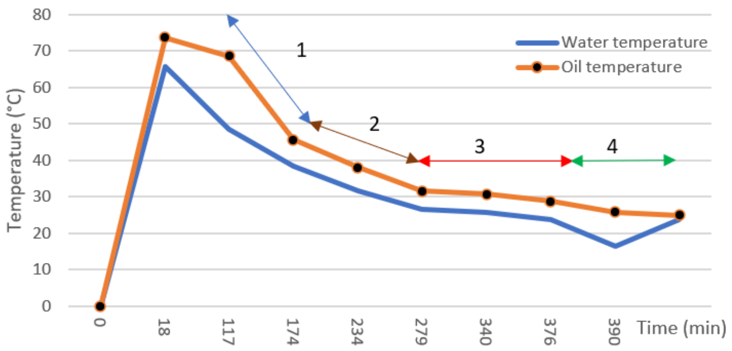

Figure 2: Theory of palm oil crystallisation

2.2 Artificial neural network

An artificial neural network utilizes the brain's nervous system to undergo data processing. If the relationship between input variables and output variables is unknown, an artificial neural network can be utilized to establish an architecture for this complex relationship. In this paper, only one hidden layer was designed. Each layer comprises multiple neurons; neurons in the same layer are not connected to each other, whereas neurons in different layer are connected to each other, forming web like patterns (Tsai et al., 2024).
The datasets were first normalised; that is, the data for each variable for the input layers were normalised in the range [-1,1]. Then, the entire data set were divided into training, validation and test data sets. The training dataset was used to reduce mean square error by adjusting the connection weights of neurons to obtain better performance. The validation data set was used to monitor the neural network for overfitting. When the mean square error of the validation data set reached the minimum, training was stopped. (Tsai et al., 2024).
To establish an optimised artificial neural network, the influence of the neural network parameters on its performance was first analysed. The discussed neural network parameters were the variables in input layer, the ratio of the size of the testing data set to the total data set, the number of neurons in the hidden layer, the training algorithm, and the initial learning rate. While investigating the influence of one parameter on the performance of the neural network, the values of other parameters were fixed.

2.3 Genetic Algorithm

Genetic algorithms are adaptive search approaches often used to solve optimisation problems. These approaches use both the equivalent of a genetic search process and the analog of natural selection. Biological nomenclature in its simplified form is kept in the genetic algorithms approach (Drissi et al., 2024).

The genetic algorithm started with a random set of individuals, known as the initial population. The selection and evolution process of genetic algorithms was designed so that the individuals with the highest fitness score were most likely to reproduce and pass on their traits to the next generation. Next, each individual was evaluated using a fitness function to calculate their fitness value. The fitness value is determined by how well the individual solution satisfies our performance objectives. This step is called individual evaluation. Then, from this population, a subset of suitable individuals was selected to create a new population for the next generation. Crossover was then carried out by crossing the selected parents (the mating pool) to produce new individuals that were better adapted. Random mutation serves as an important source of variation. In general, mutation was achieved by changing the chromosome of an individual (parent), and its probability is low. The aim of the entire procedures were that members of the new population inherit the best traits of their parents. Lastly, the selection, crossover, and mutation steps were iteratively repeated until a stopping criterion has been verified (Drissi et al., 2024).

2.4 Performance evaluation

In order to measure the effectiveness of the prediction model and evaluate neural network performance there was a main performance metrics calculated root mean square error (RMSE). RMSE is the standard deviation of the residuals (prediction errors). Residuals are a measure of how far from the regression line data points are. It is a measure of how spread out these residuals are. In other words, it explains how well the engine data is concentrated around the line of best fit. The calculation method for the RMSE is as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\tilde{y}_i - y_i)^2} \quad (1)$$

The RMSE penalizes overestimation and underestimation errors in predicting crystallisation recipe, where N is the number of samples and $\{\tilde{y}, y\}$ are predicted and desired recipe, respectively (Tsai et al., 2024).

3. Result & Discussions

In the context of CP 6 palm olein datasets, the RMSE for training, testing and validation were presented in Figure 3, indicating that the line fitted well and there is a strong correlation between predicted and actual

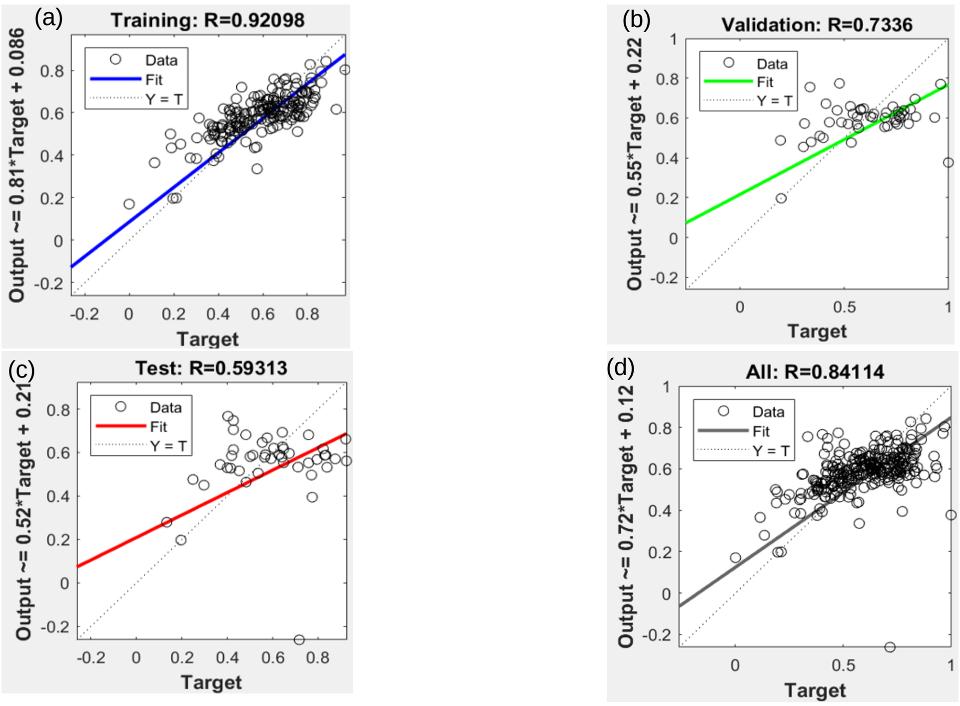

Figure 3: Linear regression for training; (b) testing; (c) validation and (d) overall of the ANN-GA in the context of CP 6 of palm olein

values (Galvanauskas et al., 2006). Besides, the relationship indicated that ANN was suitable to be used for further optimisation as it has high degree of explanatory power, robust performance beyond training set and signified that ANN maintains its predictive capability when applied to another independent datasets (Koc et al., 2007). The effectiveness of ANN to predict accurate outcome depended on how well the points fit on the line. Besides, the number of datasets could be a factor towards predictive capability of ANN. For instance, the number of datasets for CP 10 of palm olein was 50 which yielded overall RMSE of 0.4105 overall while for 400 datasets of CP 6 palm olein yielded an overall RMSE of 0.8411. The RMSE for CP 8 and CP 10 of palm olein produced were low, which can be implied that the points were not concentrated on the linear line and caused a low linear relationship between output and target values. When the number of datasets increased, the overall performance of ANN prediction capability also increased. This was because large datasets provide diverse examples, helping ANNs to generalize well to new, unseen data and reduce risk of overfitting. Thus, the large dataset was a crucial factor that determined the efficiency of ANN.

After the fitness curve was terminated, the command window would pop out the optimal inputs that will be used to run in an actual batch crystalliser for validation process as modified recipe. Table 1 was the recipe generated using ANN-GA for CP 6, 8 and 10 of palm olein after a single run.

Table 1: Optimal recipe generated from ANN-GA

Types of palm olein		CP 6	CP 8	CP 10
T-Segment No. stands for time needed to cool during the segment in minutes.	T- Segment 0	0	0	0
	T- Segment 1	15	9	9
	T- Segment 2	71	71	97
	T- Segment 3	47	55	60
	T- Segment 4	60	32	49
	T- Segment 5	40	70	40
	T- Segment 6	3	65	73
	T- Segment 7	110	58	36
	T- Segment 8	120	35	13
W-Segment No. stands for temperature of water during the segment in °C.	T- Segment 9	238		
	W-Segment 0	54	65.9	70.3
	W-Segment 1	46	46	47.5
	W-Segment 2	36	38	38.4
	W-Segment 3	27	27	31.8
	W-Segment 4	26	25	26.8
	W-Segment 5	23.5	25.2	26
	W-Segment 6	23	24.1	24
	W-Segment 7	21	16	16.7
	W-Segment 8	16	18	22.5
	W-Segment 9	14		
RBDPO IV & FFA stand for iodine value & free fatty acid of refined palm oil.	RBDPO-IV	51.5	51.2	51.8
	RDBPO-FFA	0.08	0.078	0.062
	Range of Yield (%)	52-63	66-76	76-79
	Yield (%)	63	74	77.26

For the validation process, the generated recipe from MATLAB software was utilised in SCADA system to observe the outcome.The process were carried out 3 times to prove its validity. The optimised values were within the range of lower bound and upper bound. By using the optimal recipe generated, the range of yield for CP 6 was within 52-63 % from past record. The optimal recipe gave maximum yield of 63. For CP 8 of palm olein, the yield obtained was 74 % which was within the range of yield of CP 8 of 66-76 %. The last batch was CP 10 which obtained yield of 77.26 % which fell between the range of 76-79 %. So, the poor result for CP 8 and CP 10 was due to low RMSE and it implied a low linear relationships between output and target value. Based on the result, it indicated that ANN-GA had the capability of ahieving the objective function well with the condition of large datasets to train, test and validate. In term of computational time for ANN-GA to complete a single task depend on the sample size greatly. The observation indicated that a sample size of 400 would take 15 minutes to complete.

The major contribution for this study is that the improvement of cooking oil quality provides higher nutritional value like omega-3 and omega-6, free from off-flavours, good aroma and most importantly, higher profitability from products' sales.

4. Conclusions

In short, the optimisation of palm oil started with the evaluation of overall RMSE for the performance of training, testing and validating stages. The next stage was finding the optimal parameter of the process using fitness curve in genetic algorithm. As a result, both method worked well together and successfully modeled the intricate relationships in CP 6, while CP 8 and CP 10 exhibited challenges due to reduced dataset size. Therefore, a sample size of 600-700 is recommended to improve predictive capability. The subsequent application of genetic algorithm efficiently identified optimal cooling recipes, surpassing industry standards with yields of 63 %, 74 %, and 77.26 % for CP 6, CP 8, and CP 10 palm olein, respectively. The limitations of ANN-GA approach in industrial settings, which includes varying crystallisation pathways, scalability and external resources to support data reliability. The significance of the study are the improvement in quality and energy saving in which the optimisation improves the purity of palm olein, enabling high quality of cooking oil products to be manufactured. In terms of other optimisation methods like random forest and stochastic gradient descent, ANN-GA stands out due out its ability to explore wider solution space and more robust to initial conditions as well as hyperparameters. Future research can look into multiple objective optimisation and shorter computational time to allow easier incorporation in real time process.

Nomenclature

ANN – artificial neural network
GA – genetic algorithm
IV – iodine value
CP – cloud point
RBDPO – refined, bleached, deodorized palm oil
RMSE - root mean square error
FFA- free fatty acid

Acknowledgments

The authors gratefully acknowledge the Ministry of Higher Education Malaysia (MoHE) for the financial support through Fundamental Research Grant Scheme (FRGS/1/2020/TK0/UTM/02/21).

References

Brodnitz, B., 2022, 2021/22 world palm oil stocks lowered to 12.8 MMT, Kuala Lumpur, Malaysia.
Cristea, M. V., Roman, R., Agachi, Ş. P., 2003, Neural networks based model predictive control of the drying process. In Computer Aided Chemical Engineering, Elsevier, Vol. 14, pp. 389-394.
Drissi El-Bouzaidi, M., El Bardouni, T., Chham, E., El Hajjaji, O., Nouayti, A., Idrissi, A., El-Bakkari, K., 2024, Unfolding of the neutron spectrum in a nuclear reactor using genetic algorithms method. Nuclear Engineering and Design, 422, 113156.
Galvanauskas, V., Georgieva, P., e Azevedo, S. F., 2006, Dynamic optimisation of industrial sugar crystallization process based on a hybrid (mechanistic+ ANN) model, IEEE International.
Koc, A. B., Heinemann, P. H., Ziegler, G. R., 2007, Optimization of Whole Milk Powder Processing Variables with Neural Networks and Genetic Algorithms. Food and Bioproducts Processing, 85(4), 336-343.
MacWilliams, S. V., Clulow, A. J., Gillies, G., Beattie, D. A., Krasowska, M., 2024, Recent advances in studying crystallisation of mono- and di-glycerides at oil-water interfaces. Advances in Colloid and Interface Science, 326, 103138.
Mao, J., Gao, Y., Meng, Z., 2023, Crystallization and phase behavior in mixture systems of anhydrous milk fat, palm stearin, and palm oil: Formation of eutectic crystals. Food Chemistry, 399, 133877.
Omar, Z., Hishamuddin, E., Sahri, M. M., Fauzi, S. H. M., Dian, N., Ramli, M. R., Rashid, N., 2015, Palm oil crystallisation: A review. Journal of Oil Palm Research, 27(2), 97-106.
Timms, R. E., 2005, Fractional crystallisation–the fat modification process for the 21st century. European Journal of Lipid Science and Technology, 107(1), 48-57.
Tsai, P.-H., Huang, Y., Tai, J.-H., 2024, Estimating soil water content from thermal images with an artificial neural network. CATENA, 241, 108029.
Zaliha, O., Chong, C. L., Cheow, C. S., Norizzah, A. R., Kellens, M. J., 2004, Crystallization properties of palm oil by dry fractionation. Food Chemistry, 86(2), 245-250.