Scientific Research Journal,
Vol 22, special issue (2025) 107-119
https://doi.org/10.24191/srj.v22is.13932

Yield Prediction of Supercritical Fluid Extraction of Nigella Sativa using Neutral Networks

Sarah Diana Isnin[1], Sitinoor Adeib Idris[1,2,*]

[1]School of Chemical Engineering, College of Engineering, Universiti Teknologi MARA, Shah Alam, 40450 Selangor, Malaysia
[2]Integrated Separation Technology Research Group, School of Chemical Engineering, College of Engineering, Universiti Teknologi MARA, Shah Alam, 40450, Selangor, Malaysia

[*]Corresponding Author's E-mail: adeib@uitm.edu.my

Received: 6 September 2024
Accepted: 22 October 2024
Online First: 25 January 2025

ABSTRACT

A feed-forward multi-layer neural network with Levenberg-Marquardt training algorithm was developed to predict yield for supercritical carbon dioxide (SC-CO_2) extraction of Nigella sativa essential oil. Yield of extraction depends on these variables: pressure, temperature, and extraction time hence were chosen as the input to the network. Different number of neurons in hidden layer were trained and tested using training and testing data sets. The validating data set was used to determine the network that having lowest mean-squared error (MSE) value and highest regression coefficient. The optimal ANN model, featuring four neurons in hidden layer, demonstrated high predictive accuracy with the lowest MSE of 0.42 ,1.43 and 1.25 for training, validation and test model, respectively. The regression plots indicated high R-values of 0.99641, 0.99513, and 0.98874 for the training, validation, and testing sets, respectively, confirming the model's robustness in predicting experimental data. A very good fitting between the predicted data and experimental data was observed with R^2 of 0.9891 indicates ANN shows good accuracy in predicting yield of Nigella sativa.

Keywords: Extraction; Modeling; Black Cumin Seed; Neuron; Supercritical

INTRODUCTION

Habbatus Sauda, scientifically known as *Nigella sativa*, is an herbal plant predominantly found in the Middle East. N. sativa has been traditionally used as medicine for various ailments over a long period. It is highly regarded in the Islamic community as one of the most effective means of preventing illness [1]. According to a prophetic hadith, Habbatus Sauda is a treatment for all diseases except death [2].

The oil extracted from *N. sativa* seeds is extensively utilized in the biopharmaceutical industry due to its rich phytochemical content [3]. Research indicates that *N. sativa* seeds contain significant amounts of thymoquinone (T.Q.) and thymohydroquinone [4]. T.Q. is a notable hypoallergenic bioactive component, comprising up to 30-48% of the essential oil's total composition [5]. Various studies suggest that *N. sativa* exhibits antioxidant, antihypertensive, antidiarrheal, anticancer, and anti-inflammatory properties, making it suitable for clinical applications [6]. Consequently, there is a high demand for Habbatus Sauda products due to their valuable bioactive ingredients [7]. The extraction of *N. sativa* seed oil is thus crucial in pharmaceutical manufacturing.

Traditional extraction methods for *N. sativa* seed oil involve the use of organic solvents or convection. However, these methods often result in residual solvents and low product yield. Supercritical fluid extraction (SFE) is an alternative technique that offers a safer, cleaner, and higher-yield extraction process. This method utilizes supercritical fluids (SCFs) as extracting solvents [8]. A supercritical fluid is defined as a substance at a temperature and pressure above its critical point. In this study, a Supercritical Fluid Extractor model employing supercritical carbon dioxide (SC-CO_2) was used to extract oil from N. sativa seeds. Carbon dioxide (CO_2) is considered an excellent supercritical solvent due to its environmental friendliness and Generally Recognized as Safe (GRAS) status, making it suitable for pharmaceutical and food applications [9]. Moreover, CO_2's low critical temperature is advantageous for extracting bioactive compounds from natural matrices without significant degradation [10].

Artificial Neural Networks (ANNs) are computational models inspired by the structure and function of biological neural networks. ANNs can predict outcomes by evaluating complex linear and non-linear relationships, making them superior to traditional data analysis methods in certain contexts [11,12]. ANNs offer significant advantages, including the ability to generalize from data, tolerate errors, and predict unknown test data efficiently [13]. The model is trained to minimize the error between actual and desired outputs. Once the actual output is satisfactory, the training is halted, and the weighted links between processing units are saved. The ANN model's ability to learn from historical data allows it to make accurate predictions even when faced with new, unseen conditions. This capability is particularly beneficial in optimizing extraction processes where experimental trials can be costly and time-consuming [14]. By simulating different scenarios, the ANN can identify the most efficient extraction parameters [15,16], thereby maximizing yield and reducing operational costs.

This study aims to apply ANN model to predict the extraction yield of *N. sativa* oil and to evaluate the optimal ANN architecture for extraction of *N. sativa* oil yield from SFE. Network performance was optimized by changing the number of neurons in the hidden layer, to achieve the best match of the training data with the experimental results. The results of the simulations were compared with the experimental yield data to identify the accuracy for the simulation of supercritical carbon dioxide extraction of *N. sativa*.

METHODOLOGY

Simulation Model

The data set comprises of experimental oil yield from an extraction of *N. sativa* oil using supercritical carbon dioxide extraction at different experimental conditions of pressure, temperature and time [17]. A neural network was used to predict the oil yield of *N. sativa* extracted using supercritical carbon dioxide extraction. MATLAB R2017b was employed to develop the ANN model. The parameter settings in the ANN model are listed in Table 1.

Table 1: Artificial neural network setting

Parameter	Setting
Software	Fitting app
Network algorithm	Feed-forward back-propagation
Training algorithm	Levenberg-Marquardt
Hidden layer transfer function	Logistic sigmoid-tansig
Output layer transfer function	Linear - purelin
Neuron in input layer	3
Neuron in output layer	1
Total data set	36
Performance function	Mean Squared Error

The flowchart of the training process using the fitting app is shown in Figure 1. There are three input parameters to the ANN model: temperature, pressure, and extraction time. There is one output response, which is the oil yield. The ratio of training, testing and validation are 70%, 15% and 15%, respectively.

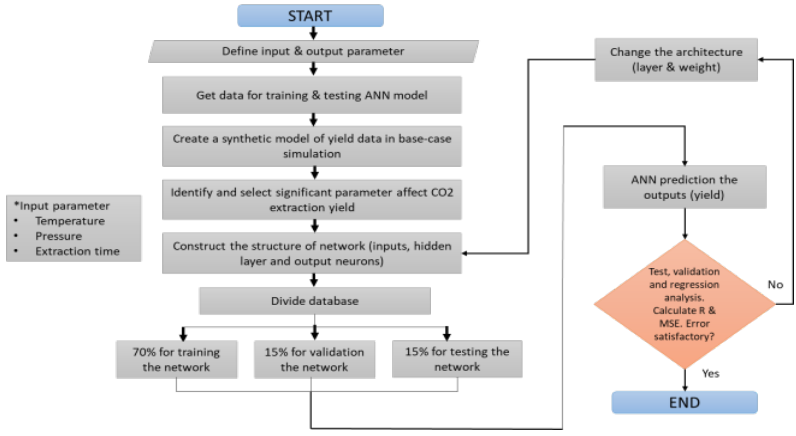

Figure 1: Flowchart of the network training process

Table 2 shows the input parameters used in this study [17]. The performance of the model can be analyzed by determining the mean square error (MSE) as shown in Eq. (1).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (1)$$

where yi is experimental data, yi is predicted data and n is number of experiment(s). The transfer function for both hidden layer and output layer were set to default. The training algorithm which is Levenberg-Marquardt was also set to default function.

Table 2: Oil yield at each parameter of extraction [17]

Pressure (bar)	Temperature (°C)	Time (min)	Experimental oil yield (g/g %)
150	40	0	0.00
		60	4.38
		90	13.14
		120	19.11
	50	0	0.00
		60	3.25
		90	9.22
		120	10.02
	60	0	0.00
		60	2.69
		90	10.11
		120	12.33
250	40	0	0.00
		60	13.05
		90	17.42
		120	21.38
	50	0	0.00
		60	11.55
		90	14.29
		120	19.98
	60	0	0.00
		60	12.88
		90	19.23
		120	21.25

350	40	0	0.00
		60	14.25
		90	15.68
		120	22.33
	50	0	0.00
		60	11.25
		90	16.49
		120	17.92
	60	0	0.00
		60	12.56
		90	18.31
		120	23.20

Model Training Procedure

One input layer including extraction time, pressure, and temperature, one output layer with neuron representing yield, and one hidden layer with associated inputs and output. A neuron is a processing element that takes several inputs as well as their weights, sums them up, adds a bias (b) and uses the results as an argument for a transfer function (f). A transfer function is assigned to each neuron that determines the value of the outputs. The network is trained by modifying the weights through minimizing the errors. Introducing new weights is continued until the outputs meet the desired values.

RESULTS AND DISCUSSION

The Network Optimization of the ANN Model

Ten configurations (one hidden layer with one until ten neurons) of the ANN model have been tried out to establish the optimum network for the best prediction performance. Table 3 lists the distribution of dataset allocated to training (70%), validation (15%), and testing (15%) for determining the optimal number of hidden neurons in the neural network.

Table 3: Distribution of dataset in ANN group

Group	Dataset
Training	0, 4.38, 19.11, 3.25, 10.02, 0, 2.69, 10.11, 12.33, 0, 13.05, 17.42, 21.38, 11.55, 14.29, 19.98, 21.25, 0, 14.25, 15.68, 22.33, 0, 11.25, 16.49,18.31 & 23.2
Validation	12.88, 19.23, 17.92, 0 & 12.56
Testing	13.14, 0, 9.22, 0, & 0

Good fitting between the experimental data and the data predicted from the neural network when applying the 70:15:15 ratio in predicting oil yield from plant parts as what reported by previous study [8,13]. As shown in Figure 2, the mean squared error (MSE) values for all models (training, validation, and test) were evaluated across different numbers of neurons in the hidden layer [8]. The ANN prediction was stopped at 10 neurons because the MSE for the test model began to increase, deviating from the objective of achieving an MSE close to zero [18]. After training, it was observed that network having one hidden layer with four neurons gave minimum errors with MSE values of 0.42, 1.43, and 1.25 for the training, validation, and test models, respectively as depicted in Figure 2.

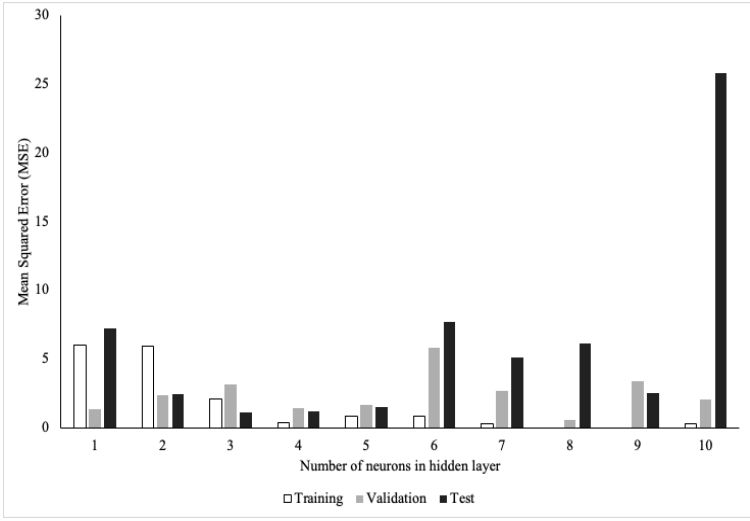

Figure 2: Mean Squared Error for each model at each number of neurons in hidden layer.

Performance of the Developed ANN Model

Figure 3 illustrates the optimal ANN structure 3-4-1, which denotes the number of neurons in the input, hidden, and output layers, respectively. It also indicates the optimal reliability of the ANN model developed for N. sativa oil extracted using SFE.

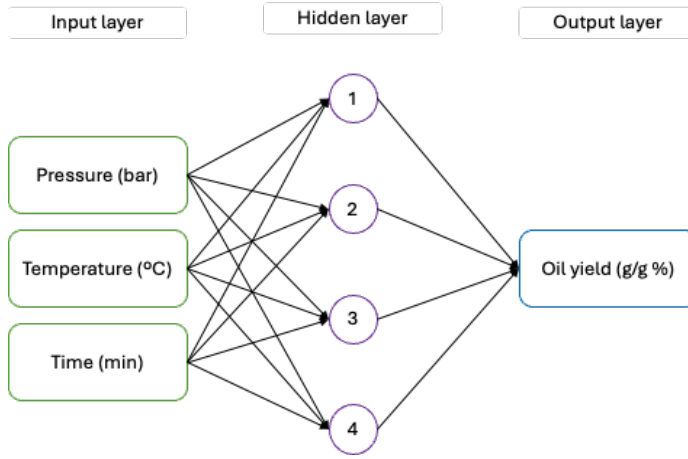

Figure 3: Optimum architecture of Artificial Neural Network for *N. sativa* oil yield prediction (4 hidden neurons)

Figure 4 shows the regression plot comparing the target values with the predicted values for each model. The correlation coefficients for training, validation, and testing are close to 1 with a regression coefficient (R^2) of 99.5%, indicating a high degree of fit between the experimental and predicted oil yield given the pressure, temperature and time of extraction. In this study, almost all data points lie along the 45° line in the regression plots (Figure 4), demonstrating the compatibility of the experimental data with the predictions made by the ANN [19]. Figure 5 shows the data from experiment and data predicted from ANN. A very good fitting between the predicted data and experimental data was observed from the figure with R^2 of 0.9891. The figure show that the predicting model agree with the experimental data. A value of R^2 close to 1 is crucial for selecting the best number of neurons in the ANN hidden layer [20]. Therefore, ANN as the model to predict the oil yield data from *N. sativa* was suitable and was supported and reported by several previous study [21].

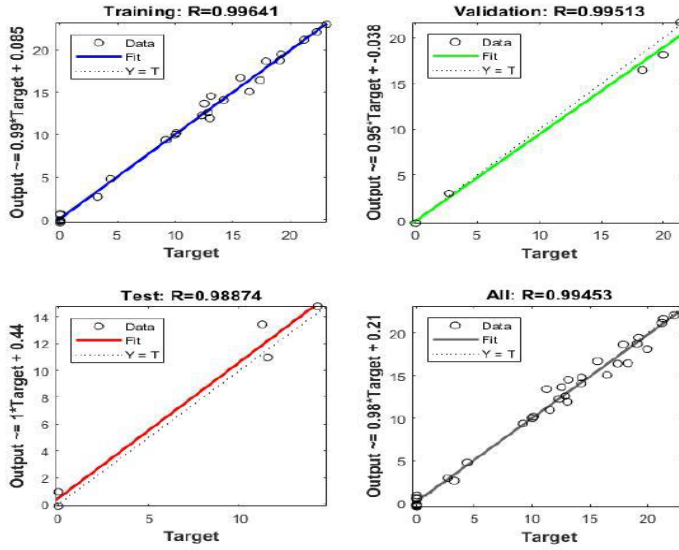

Figure 4: Regression plot for training, validation, test and all with R values

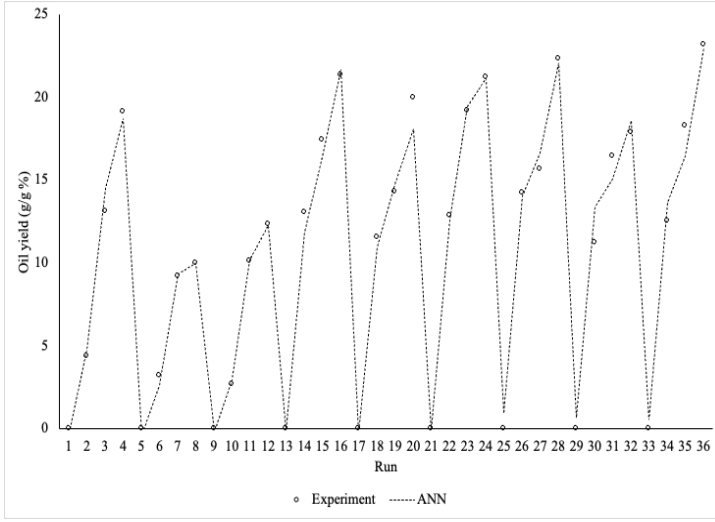

Figure 5: Comparison between experimental oil yield and predicted oil yield using ANN at each run

CONCLUSION

ANN has successfully been used to predict the oil yield data of *N. sativa* with the optimized configuration of 3-4-1. When the number of neurons in the hidden layer was four, the regression value for training and testing were 0.99641 and 0.98874, respectively. The MSE values were 0.42 and 1.45 for training and testing, respectively. The comparison between the experimental oil yield and the predicted yield for each run shows the best fitting with R^2 value of 0.9891 and R^2 of above 0.98 for all model in the regression plot. In this research, the predictive model, ANN is very compatible in predicting yield from N. sativa oil extraction using supercritical carbon dioxide.

ACKNOWLEDGEMENT

The authors gratefully acknowledge College of Engineering, Universiti Teknologi MARA (UiTM) for the financial and other support to this research. This research was not funded by any grant.

REFERENCES

[1] N. Y. Santiworakun, A. Suksuwan, S. Sirikwanpong, W. Dahlan and T. Ariyapitipun, 2022. Physicochemical characterization of microcapsules containing cold pressed black cumin seed oils (Nigella sativa L.) as an alternative nutrient source in a functional diet, *Lebensmittel-Wissenschaft & Technologie-Food Science and Tenchnology,* 157, 1-8.

[2] S.A. Abd Ghafar, M. F. Yakop, R. Mohamad Hanafiah and N. Ismail, 2023. Chapter Thirteen - Antihypercholesterolemic and antihyperglycemic activity of Nigella sativa, Biochemistry, Nutrition, and Therapeutics of Black Cumin Seed (A.A. Mariod, ed), *Academic Press.* 143–159.

[3] S. Dimitrijević, M. Milić, V. Tadić, S. Maksimović, V. Filipović, S. Dimitrijević-Branković, M. Miljković and I. Salamon, 2024. Black cumin essential oil as a valuable source of bioactive compounds:

Evaluation of the conventional vs. modern extraction technique, *Sustainable Chemistry and Pharmacy,* 37, 1-10.

[4] W. Sakdasri, P. Sila-ngam, S. Chummengyen, A. Sukruay, S. Ngamprasertsith, W. Supang and R. Sawangkeaw, 2023. Optimization of yield and thymoquinone content of screw press-extracted black cumin seed oil using response surface methodology, *Industrial Crops and Products,* 191, 1-7.

[5] S. Veeramani, A. P. Narayanan, K. Yuvaraj, R. Sivaramakrishnan, A. Pugazhendhi, I. Rishivarathan I., S. P. Jose and R. Ilangovan, 2022. Nigella sativa flavonoids surface coated gold NPs (Au-NPs) enhancing antioxidant and anti-diabetic activity, *Process Biochemistry,* 114, 193–202.

[6] I. M. Rizwanul Fattah, R. S. M. Ashrafur and A. Ahmed, 2023. Chapter Eight - Black cumin (Nigella sativa L.) essential oil and aroma quality. Biochemistry, Nutrition, and Therapeutics of Black Cumin Seed (A.A. Mariod, ed), *Academic Press.* 71–87.

[7] V. Manjunatha, J. E. Nixon, G. F. Mathis, B. S. Lumpkins, Z. B. Güzel-Seydim, A. C. Seydim, A. K. Greene and X. Jiang, 2023. Nigella sativa as an antibiotic alternative to promote growth and enhance health of broilers challenged with Eimeria maxima and Clostridium perfringens, *Poultry Science,* 102 (8), 1-12.

[8] S. A. Idris, M. Markom, N. Abd Rahman and J. Mohd Ali, 2022. Prediction of overall yield of Gynura procumbens from ethanol-water + supercritical CO2 extraction using artificial neural network model, *Case Studies in Chemical and Environmental Engineering*, 5, 100175, 1-9.

[9] F. Khodaie and S. M. Ghoreishi, 2021. Experimental extraction of gallic acid from brown sumac seed (Rhus coriaria) using supercritical carbon dioxide and ethanol as co-solvent: Modeling and optimization, *Journal of Supercritical Fluids,* 175 (April), 105266.

[10] N. Hassim, M. Markom, M. I. Rosli and S. Harun, 2021. Scale-up approach for supercritical fluid extraction with ethanol–water modified carbon dioxide on Phyllanthus niruri for safe enriched herbal extracts, *Scientific Reports*, 11 (1), 1–19.

[11] J. Chu, X. Liu, Z. Zhang, Y. Zhang and M. He, 2021. A novel method overcomeing overfitting of artificial neural network for accurate prediction: Application on thermophysical property of natural gas, *Case Studies in Thermal Engineering,* 28 (July), 101406, 1-13.

[12] J. Li, Q. Li, H. Hao and L. Li, 2021. Prediction of BLEVE blast loading using CFD and artificial neural network, *Process Safety and Environmental Protection,* 149, 711–723.

[13] D. Ramírez-Brewer, S. E. Quintana and L. A. García-Zapateiro, 2024. Modeling and optimization of microwave-assisted extraction of total phenolics content from mango (Mangifera indica) peel using response surface methodology (RSM) and artificial neural networks (ANN), *Food Chemistry X*, 22, 1-11.

[14] S. Rajasekaran and G. A. V. Pai, 2017. Neural Networks, Fuzzy Systems and Evolutionary Algorithms : Synthesis and Applications. Prentice-Hall of India. Available at: https://books.google.com.my/books?id=D1XhDgAAQBAJ.

[15] J. L. Pilkington, C. Preston and R. L. Gomes, 2014. Comparison of response surface methodology (RSM) and artificial neural networks (ANN) towards efficient extraction of artemisinin from Artemisia annua, *Industrial Crops and Products,* 58, 15–24.

[16] B. Suryawanshi and B. Mohanty, 2018. Industrial Crops & Products Application of an artificial neural network model for the supercritical fluid extraction of seed oil from Argemone mexicana L seeds. *Industrial Crops and Products,* 123 (June), 64–74.

[17] Z. Solati, B. S. Baharin and H. Bagheri, 2012. Supercritical carbon dioxide (SC-CO2) extraction of Nigella sativa L. oil using full factorial

design, *Industrial Crops and Products,* 36 (1), 519–523.

[18] B. Pavlic, O. Bera, S. Vidovic, L. Ilic and Z. Zekovic, 2017. Extraction kinetics and ANN simulation of supercritical fluid extraction of sage herbal dust, *Journal of Supercritical Fluids,* 130, 327–336.

[19] Z. Zeković, O. Bera, S. Đurović and B. Pavlić, 2017. Supercritical fluid extraction of coriander seeds: Kinetics modelling and ANN optimization, *Journal of Supercritical Fluids*, 125, 88–95.

[20] D. Graupe, 2019. Principles Of Artificial Neural Networks: Basic Designs to Deep Learning (4th Edition). World Scientific Publishing Company. Available at: https://books.google.com.my/books?id=77uSDwAAQBAJ.

[21] G. Sodeifian, S. A. Sajadian and N. S. Ardestani N.S, 2016. Evaluation of the response surface and hybrid artificial neural network-genetic algorithm methodologies to determine extraction yield of Ferulago angulata through supercritical fluid, *Journal of Taiwan Institute of Chemical Engineers,* 60, 165–173.