CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Artificial neural network–genetic algorithm-based optimization of biodiesel production from Simarouba glauca
Authors	Sivamani, Selvaraju;Selvakumar, Selvaraj;Rajendran, Karthik;Muthusamy, Shanmugaprakash
Publication date	2018-02-13
Original Citation	Sivamani, S., Selvakumar, S., Rajendran, K. and Muthusamy, S. (2018) 'Artificial neural network–genetic algorithm-based optimization of biodiesel production from Simarouba glauca', Biofuels, In Press, doi:10.1080/17597269.2018.1432267
Type of publication	Article (peer-reviewed)
Link to publisher's version	10.1080/17597269.2018.1432267
Rights	© 2018. This is an Accepted Manuscript of an article published by Taylor & Francis in Biofuels on 13 Feb 2018, available online: http://www.tandfonline.com/10.1080/17597269.2018.1432267
Download date	2026-10-30 18:04:37
Item downloaded from	https://hdl.handle.net/10468/5620



UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Artificial neural network-genetic algorithm based optimization of biodiesel production from *Simarouba glauca*

Selvaraju Sivamani[1], Selvaraj Selvakumar[1], Karthik Rajendran[2,3], Muthusamy Shanmugaprakash[1,*]

[1]Downstream processing laboratory, Department of Biotechnology, Kumaraguru College of Technology, Coimbatore641049, India.

[2]Department of Biological and Ecological Engineering, Oregon State University, Corvallis, United States

[3]MaREI Center, Environmental Research Institute, University College Cork, Cork, Ireland

*Corresponding author

sunbionin@gmail.com

+91 99423 11658

Abstract

Transesterification reaction was carried out employing anoil of paradise kernel (*Simarouba glauca*), a non-edible source for producing *Simarouba glauca* methyl ester (SGME) or biodiesel. In this study undertaken, the effect of 3 variables such as reaction temperature, oil to alcohol ratio and reaction time were studied and optimized using response surface methodology (RSM) and artificial neural network (ANN) on free fatty acid (FFA) level. Formation of methyl esters due to areduction in FFA are observed in gas chromatography-mass spectroscopic (GC-MS) analysis. It was inferred that optimum conditions such as an oil to alcohol ratio of 1:6.22, thetemperature of 67.25 and time duration of 20 h produce abetter yield of biodiesel with FFA of 0.765±0.92%. The fuel properties of paradise oil meet the requirements for biodiesel, by Indian standards. The results indicate that the model is in substantial agreement with the research and Simarouba oil can be considered a potential oil source for biodiesel production.

Keywords: biodiesel; optimization; transesterification; artificial neural network; response surface methodology

Introduction

Self-reliance on energy resources is vital for economic development in India and other developing countries. The need to search for alternative sources of energy that are renewable, safe and pollution-free is prioritized, in view that fossil fuel sources are finite and pose deteriorating effects on the atmosphere through the release of sequestered carbon such as CO_2 and CO, thereby inflicting global warming. In addition, the uncertainty of supplies and drastic price hikes of fossil fuels in the international market are posing serious threats to the economy of developing countries [1]. To overcome this threat, different forms of renewable energy has been exploited including solar energy, wind mills, hydrothermal and bioenergy. As India is primarily an agricultural country, where more than 50% of the land is cultivable, it is essential to promote bioenergy based systems. In 2012, biodiesel production in India was 268-m^3/day in 2012 which is equivalent to 0.06% share of the total petroleum consumed during the same period [2] . The major plant sources for biodiesel production in India are*Jatropha curcas, Pongamia pinnata, Azadirachta indica, Simarouba glauca, Madhuca latifolia, Salvadora persica, Garcinia indica, and Simmondsia chinensis.*Among them, *Simarouba glauca*, a polygamodioecious perennial popularized as an important oilseed tree crop for cultivation on wasteland and native to Central America, popularly known as aceituno was introduced to India in the early 1960s[3]. It is a perennial tree with a profusely green and glossy foliage on its rounded crown and has the inherent nature of growing branches that sprawl the ground levels since early stages of growth. The plant begins fruition right from 5-6 years of age and produces yield until 50-60 years. A mature plant can yield up to 20-50 kg of brown-coloured oval-shaped seeds [4]. It has been reported that the seeds of the above plant contain 40% kernels and the kernels contain 60% fat, which is edible. The average values of major fatty acids present in paradise oil were as follows:

oleic (51.1%), stearic (30.4%), and palmitic (12.6%) [5], [6]. Furthermore, it has been reported that the characteristics of the fat and the fatty acid profile of Indian origin varieties do not significantly differ from those reports of seeds from other countries [7] [8].

Various researchers have attempted to extract biodiesel from different plant sources. For instance, pradhan et. al [9] investigated the oil content of tree borne oilseeds such as jatropha, karanja and simarouba kernel. In this study, it was found that simarouba kernel has highest oil content (61.04%) when compared to *jatropha* (45.03%) and karanja (34.67%) kernel. Dash et al. [10] examined the physical properties of *Simarouba glauca.L* kernel and found that it had 8.51% moisture and 65.8% oil content. The average length, width, thickness and 1000 unit mass for kernel were 13.78 mm, 7.77 mm, 6.71 mm and 330.26 g respectively. Bulk density, true density, porosity and angle of repose of kernel were 727.73 kg/m^3, 1019.3 kg/m^3, 28.61% and 35.02°C correspondingly. Duhan et al. (2011) observed that the refining of simarouba oil resulted in the improvement of oil quality, thereby making the oil more suitable for human consumption and biodiesel production. Refinement of oil is aimed at removing the impurities to make oil adaptable for edible and biodiesel purposes.

Devanand Mahalakshmi [11] carried out engine tests with the objective of obtaining emission and combustion characteristics of a diesel engine running on methyl ester of paradise oil (MEPO) and its diesel blends. From the engine analysis, it was found that the performance of MEPO and its diesel blends were similar to that of standard diesel. Devan and Mahalakshmi [11, 12] explored the replacement of diesel fuel with biodiesel, namely methyl esters of paradise and eucalyptus oils. Various proportions of paradise oil and eucalyptus oil were prepared by volume basis and a methyl ester derived from paradise oil is considered an ignition improver. The

combustion characteristics of Me50–Eu50 blend were comparable to those of diesel. Jena et al. (2010) demonstrated a suitable process comprising acid pre-treatment followed by base transesterification reaction to produce biodiesel from a mixture of Mahua and *Simarouba* oils with high FFA.

Multi-variant techniques are being widely used for the past few years for the purpose of process optimization [13]. Most of the previous research on biodiesel extraction usually involve the effect of one-variable at a time (OVAT) on the experimental output, while other parameters were maintained constant; this is not only time-consuming but expensive as well in the consumption of reagents and materials [14]. Moreover, this approach does not provide any information about the effect of interactions effect among the different variables involved. It is based on response surface methodology that can be applied for process optimization through the development of the mathematical model. Central composite and Box-Behnken designs are among the principal response surface methodologies used in experimental design [15]. Using Box-Behnken design of experiments, a mathematical correlation between alcohol to oil molar ratio, catalyst concentration and temperature were determined to obtain optimal biodiesel yield [16].

However, RSM based models are exact for only a limited range of input process parameters, and thus, impose a limitation on the use of RSM models which are highly non-linear processes such as biosorption [17, 18, 19].These constraints have led to the use of an artificial neural network (ANN) for developing an empirical model for a non-linear system [20] [21]. The ability of an ANN to learn and generalize the behavior of any complex and non-linear process makes it a powerful modeling tool [22]. In recent years, response surface methodology (RSM) and artificial neural network (ANN) methods have been used concurrently for both modeling and optimization

of a various engineering applications [23],[24],[25]. However, both ANN and RSM based models have different edges over the other in the predictive and optimization capacities of a given system, instigating researchers to compare the results of the different approaches, to understand the processes under investigation even better[22].

Therefore in the present study, 3-level three factors Box-Behnken and ANN-GA based models have been developed to predict the relationship between the input: methanol to oil ratio (M), reaction temperature (T) and reaction time (t) and output variables: FFA level of crude *simarouba* oil. Subsequently, the results predicted by the ANN and RSM techniques were compared statistically to the coefficient of determination (R^2), root mean square error (RMSE) and absolute average deviation (AAD) based on the validation data set for their predictive and generalization capabilities. A genetic algorithm coupled with an artificial neural network (GA-ANN) was used to obtain the best pre-treatment process parameters to reduce FFA levels to around 1%. The quality of biodiesel was analyzed by gas chromatography (GC) and the fuel properties of SGME were compared to EN 14214 and ASTM D6751 biodiesel standards. To our knowledge, this is the first detailed analysis available for biodiesel production from *simaruba* seeds by coupling both RSM and ANN methodology.

Materials and Methods

Production of biodiesel from simarouba seeds

Simarouba seeds were procured from Forest College and Research Institute, Mettupalayam, Coimbatore district, Tamil Nadu, India. Methanol with a purity of 99.5% and Potassium Hydroxide (KOH) were purchased from Merck Company. Simarouba oil was extracted from the kernel as intended. The batch reaction kinetic experiments were employed to optimize alcohol to

oil ratio, temperature and time in the production of SGME. The trans-esterification reactions are performed in various conditions to determine the optimum conditions of trans-esterification. 100 mL of simarouba oil was poured in to the reactor and allowed to equilibrate to the temperature of reaction at 300 rpm. Hot water circulated through the jacket of the reactor provided the necessary heat for the reaction. 0.7% (w/v) of KOH catalyst was dissolved in varying volume of methanol. On attaining the required temperature, the catalyst was added to the reactant and was maintained for 1 h until completion of the reaction. After 1 h, the trans-esterification reaction was completed and mixture was withdrawn from the reactor and poured in the separating funnel to separate SGME from glycerol. Gravity separation of two phases requires at least 1 h. Glycerol and SGME have deep red and bright yellow colour respectively. After separation of biodiesel, it should be washed to remove impurities and left out reactants. The biodiesel was washed out 10 times. Initially, the washing process should be done slowly and carefully to avoid soap formation. One litre of warm distilled water was used per litre of biodiesel. The washing procedure should be repeated until the color of water transforms to white. Finally, biodiesel was dried completely using silica gel.

Experimental design

Response surface methodology

In the study, the effects of the operating parameters (viz., methanol to oil ratio (M), reaction temperature (T) and reaction time (t) and output variables (FFA level of crude simarouba oil) were evaluated using response surface methodology (RSM). Response surface methodology (RSM) is an experimental technique which consists of mathematical and statistical methods for

designing experiments, building models and evaluating the effects of the variables, used to measure the optimal response within the specified range [26]. The coded and non-coded levels of the independent variables are given in Table 1. Two replications were carried out for all design points except the centre point (0,0,0), and the experiments were carried out in a randomized order. It lead to coefficients for each term present in the model, and these values were used to represent the relationship between the dependent and independent variables used in the experiments.

The response variable: Free fatty acid level of crude simarouba oil (%) can be expressed as a function of the independent variables, according to the following equation:

$$q_e(mg/g) = \beta_o + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{1 \le i \le j}^{k} \beta_{ij} x_i x_j + \varepsilon \qquad (1)$$

Where β_o is the constant coefficient, β_i, β_{ii} and β_{ij} are the coefficients for the linear, quadratic and interaction effects respectively, x_i and x_j are the independent variables and ε is the error. In this study, the statistical and graphical software package, Design Expert Trail 8.0.7.1 (Stat-Ease, USA) was used for the regression analysis, graphical analysis and analysis of variance (ANOVA).

Artificial neural network

Artificial neural networks are computer models based on a simplified modeling of the brain's biological functions, viz., having the ability to learn, think, remember reasons and solve problems. Conceptually, neural network models are composed of neurons and weights, and are based on the principle that a highly interconnected system of simple processing elements can learn complex interrelationships between independent and dependent variables [21]. Artificial

neural networks (ANN) were introduced recently into the field of engineering studies as a tool for modeling and optimization of systematic parameter studies involved in the particular process. ANNs possessing a highly interconnected structure, consist of a large number of simple processing elements called neurons which are arranged in different layers in the networks: an input layer(s), an output layer(s) and all other intermediate units called hidden layers. Each of these layers consists of a number of inter-connected processing units called neurons. These neurons are interrelated to one another by exchanging the signals via adjusted weighted connections. The input layer receives signals from external sources, is weighted individually and sends this information for processing to the hidden layers [27]. The hidden layer then does all the pre-processing and gives the output based on the sum of the weighted values from the input layer, modified by a sigmoid transfer function (transig) at the hidden layer and a linear transfer function (purelin) as output were used. The most popular ANN multilayer perception (MLP) is a supervised learning technique that compares the responses of the output units to the desired responses, and adjusts the weights in the network so that the next time when the same input data is presented to the network, the network' s response will be very much closer to the desired response. The most important feature is that the learning process of the neural network and the information obtained are used to store across the network weights. The network is trained to make proper associations between the inputs and the respective outputs [28, 29].

The performance of the ANN process can be expressed in terms of the root mean square error (*RMSE*) and correlation coefficients (R^2), and is given by

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}\left(X_{im} - X_{ip}\right)^2} \qquad (2)$$

Where n is the number of data points, and X_{im} and X_{ip} are the measured and predicted values of the processes respectively. The final network was selected based on the lowest error in the train and depending upon the test data. The input and output data to the individual ANN nodes were normalized within a range of -1 to 1 in order to achieve fast convergence to obtain the minimal RMSE values. The output values obtained from the ANN are also in the range of -1 to 1, and converted to their original data based on the reverse method of normalization.

The normalized values of each input and output are obtained according to the following formula [18]

$$Normalised = \left[\frac{2*(X_{Ac} - X_{\min})}{(X_{\max} - X_{\min})}\right] - 1 \quad (3)$$

Where $X_{\min}$, $X_{\max}$ and X_{Ac} are the minimum, maximum and actual data respectively.

For this purpose, the Neural network Toolbox of MATLAB mathematical software (MATLAB 7.5.0.0.342 (R2007 b), (The Mathworks Inc., MA,USA) was utilised throughout the study.

Results and Discussion

Response surface model

Experimental values obtained for FFA level responses at the design points are given in Table 2. Multiple regression coefficients as indicated in Table 3 were obtained by employing a least square technique to predict quadratic polynomial model for the FFA level. Analysis of these parameters with the t-test indicated that all linear and quadratic terms had significant effect on

the reduction of FFA level. The interaction terms were also found to be significant. The model was tested for adequacy by analysis of variance. The regression model for data was found to be highly significant with a coefficient of determination as 0.9799. Using the coefficients determined, the predicted model for FFA level is:

$$FFA = -17.32 - 0.0.609*M - 0.43*T - 0.18*t - 0025*M*T - 0.000651*M*t - 0.0009*T*t + 0.059M^2 + 0.0036\,T^2 + 0.007*t^2 \quad (4)$$

where FFA is the Free fatty acid, M,T and t are the coded values of the given variables.

From Equation (4), it could be concluded that the linear and quadratic effects of M and t were the primary determining factors of the responses followed by quadratic effect. The relationship between independent and dependent variables of the developed model is shown in Figure 1 (a-c) in the form of surface plots. All Eigen values obtained for the analysis were positive indicating that the stationary point for the response at minimum.

The responses corresponding to the second-order model indicated that as feed temperature increases, FFA conversion increases with increasing methanol to oil ratio and reaction time (Fig. 1 a). Maximum conversions were therefore obtained for low volume of methanol, due to the fact that it was the most significant factor with positive effect. However, at higher feed temperature, there seemed to be an amplified effect of increasing methanol amount or reaction time. This could be due to the negative coefficients of methanol-time interactions[30]. As more and more FFA was trans-esterified by methanol, the reaction might have produced more water with increasing time. It was observed that high amount of methanol

(methanol to oil ratio of 1:3) was required to reduce the FFA level below 1% (Fig. 1b). However, in a commercial application, this could be reduced by recovering part of it by fractional distillation of the water mixture which would be separated from the top. Also it could be possible to further reduce the methanol consumption by continuously removing the water from the mixture during the reaction. At low feed temperature, there was a moderate increase in FFA level with reaction time, since the time effect was negative and greater than that of feed temperature. For higher feed temperature the increase of conversion with time became lower as a result of the positive interaction. Therefore, the operating conditions of high feed temperature and reaction time favoured the biodiesel production. It was also observed that increasing reaction time beyond 24 hours does not have much effect on reducing the FFA level as desired (Fig. 1c). This might be due to the effect of water produced during the esterification of FFA, which prevented further reaction. Verification experiments showed reasonably close value of 0.438 ± 0.92% to the predicted value for the stationary point (0.417%) which was in accordance with the predicted RSM model.

Artificial neural network

Using the experimental results in Table 2, ANN was also applied to predict the relationship between FFA and the three process variables. In this study, different training algorithms were tested by varying the number of hidden layers and neurons by training the different feed-forward networks of various topologies, in order to select the optimal architecture based on the minimization of the performance function – the Mean square error (MSE). Table 7 shows the best combination of the ANN parameters that were able to predict the output parameter more accurately. The learning rate and the error goal were selected, based on trial and error, in order to keep the minimum distance between the experimental and predicted values [31]. The optimum

ANN architecture taken as 3-8-1 (three neurons in the input layer, eight in the hidden layer and one in the output layer) were chosen by trial and error, when the mean square error (MSE) decreased gradually and became constant. The values of MSE obtained from the ANNs for both the batch and continuous modes were 0.00458 which are close to the acceptance limit for the MSE set to zero. The closeness of the training and testing errors validates the accuracy of the model. The coefficient of determination: R-squared (R^2) indicates the goodness of fit between the experimental and predicted responses given by the ANN model. In this case, the values of R^2 for the model is nearly equal to 1, indicating the significance of the model (Fig. 2). Thus good, sufficient and higher values of coefficient (R^2) of the ANN predicted and experimental responses for biodiesel production suggest that the newly constructed ANN has the ability to predict the optimal conditions for biodiesel production. The weight and threshold values of each layer, which determined the structure of the trained ANN, were as follow and given in Equation (5):

$$net.iw\{1\} = \begin{pmatrix} -1.754 & 1.32 & 1.86 \\ 1.09 & -1.3 & -2.47 \\ -1.6 & 2.14 & -0.05 \\ -1.4 & -2.4 & -0.73 \end{pmatrix}$$

$$net.iw\{2\} = (0.18 \quad 0.53 \quad -0.21 \quad -0.71 \quad 0.49 \quad 0.48 \quad 0.40)$$

$$net.b\{1\} = (2.5 \quad 1.38 \quad 1.13 \quad 1.42 \quad 0.76 \quad 1.92 \quad 3.54)^T$$

$$net.b\{2\} = (0.504)^T$$

(5)

Comparison of RSM and ANN

After RSM and ANN models were built, the predictive capability of both the ANN and RSM models were compared. A completely new set of experiments was conducted which does not belong to the training data sets. The experimental and predicted values of the responses along with their residual values for both the ANN and RSM models are given in Table 4 and visually represented in Fig 3a, b. Also, the prediction abilities of the newly constructed ANN and RSM models were statistically measured, in terms of the root mean square error (RMSE), co-efficient of determination (R^2) and absolute average deviation (AAD) [22] and given in Table 5 . The predictive values obtained from the RSM and ANN models for FFA was found to be as significant as those obtained using experiments according to the DOE. This proves the applicability of the ANN and RSM in the prediction and optimization of the biodiesel production with the minimum experimental set-up which minimizes the consumption of reagents, which in turn reduces the effluent treatment cost. Also, from Table 5, it was found that predicted values by ANN model were much closer to experimentally measured ones hinting that ANN model was quite successful for both simulation and predicted values. Similar observations were obtained by many research groups to study various engineering problems. [32] ;[33].

Optimization of biodiesel production using ANN-GA

GA is a class of parallel iterative and global search algorithm with certain learning abilities employing crossover and mutation operators to solve optimization problems [34]. In order to obtain the global optimal solution, an attempt has been made using GA for minimal FFA condition. In order to achieve that, the output of the developed ANN model was adopted in

calculating the values of the fitness function for the GA routine available in MATLAB Genetic Algorithm toolbox. The best fitness plot for the GA maps the gradual convergence of the best fitness values of successive generations towards the final optimum value as shown in Fig. 4. Values of the best combination of biodiesel production using Alcohol ratio, temperature and time is 1:6.22, 67.25 and 9.4 h , respectively.

Characterization of biodiesel produced from optimal condition

A test case of alkali catalyzed trans-esterification was run under optimal condition predicted by ANN coupled with GA using simarouba oil that had FFA level of 1.683%. A 6.22 molar ratio of methanol to oil molar ratio and an alkali catalyst (0.7% w/v KOH) were used for the reaction. Following this process, a yield of 62% biodiesel was obtained from simarouba oil. The fuel properties of SGME obtained are summarized in Table 6 along with crude simarouba oil, karanja and jatropha biodiesel and diesel. It can be seen that the fuel properties of simarouba biodiesel are similar to those of karanja and jatropha biodiesel except dynamic viscosity, kinematic viscosity and FFA level. Moreover, simarouba biodiesel had comparable fuel properties with those of diesel and conforming to the latest Indian standards for biodiesel [35] ;[36].

The GC-MS chromatogram for SGME obtained from crude simarouba oil is shown in Figure 5(a). From this spectrum, it was found that the amount of oleic methyl ester was 55.74% in biodiesel followed by methyl stearate and palmitate. The m/z for methyl palmitate, methyl oleate, methyl stearate and methyl linoleate was observed to be 74, 87, 143.1 and 171.1, respectively. The high peak shown in Fig. 5b, 6 at m/z of 111.1 is n-hexane and should not be

confused with other methyl ester peaks. The overall conversion of oil to methyl esters was calculated on the basis of their relative area percentage and was found to be above 62%.

Conclusion

In this study, the modeling, predictive and generalization capabilities of RSM and ANN models were compared for biodiesel production using simaruba seeds. The BBD based and a feed-forward multilayered perceptron (MLP) ANN trained models were developed to predict the FFA (%). From the experimental data, it was inferred that the simulation and prediction of ANN were superior to that of RSM model and the performance of two models were statistically confirmed. To conclude, the model developed by ANN has superior ability than the RSM model due its global random search and high prediction accuracy in biological systems.

Acknowledgements

The authors are thankful to the management of Kumaraguru College of Technology, Coimbatore, India, for the providing the research facilities.

References

1. Wani S, Sreedevi T. Pongamia's journey from forest to micro-enterprise for improving livelihoods. Pongamia as a source of Biodiesel in India. 2005;189.
2. Karmakar A, Karmakar S, Mukherjee S. Biodiesel production from neem towards feedstock diversification: Indian perspective. Renewable and Sustainable Energy Reviews. 2012;16(1):1050-1060.
3. . Available from: http://www.eia.gov/beta/international/index.cfm

4. Pradhan K. Commercial exploitation of Simarouba glauca. Strategies for development of tree borne oil seeds and niger National Oilseeds and Vegetable Oils Development (NOVOD) board, Ministry of Agriculture, Govt of India. 1995:88-100.

5. Coleman M. The pancreatic hydrolysis of natural fats. IV. Some exotic seed fats. Journal of the American Oil Chemists Society. 1965;42(9):751-754.

6. Chikara J, Shethia B, Rathod M, et al. Simarouba glauca-a rich-source of oleic acid for rehabilitation of marginal lands of the country. JOURNAL-OIL TECHNOLOGISTS ASSOCIATION OF INDIA. 1998;30:177-180.

7. Rao K, Lakshminarayana G. CHARACTERISTICS AND COMPOSITION OF SIMAROUBA GLAUCO DC SEEDS AND FAT. Journal of Food Science and Technology. 1983;20.

8. Rath S, Srinivasulu C, Mahapatra S. Investigations on Simarouba glauca—a new oilseed of Indian origin. Journal Oil Technologists Association of India. 1987;19:64-65.

9. Pradhan R, Naik S, Bhatnagar N, et al. Physical Properties of Tree Borne Oil Seed: Potential Biodiesel Feedstock in India. Assam University Journal of Science and Technology. 2010;4(2):46-53.

10. Dash A, Pradhan R, Das L, et al. Some physical properties of simarouba fruit and kernel. International Agrophysics. 2008;22(2):111.

11. Devan P, Mahalakshmi N. Utilization of unattended methyl ester of paradise oil as fuel in diesel engine. Fuel. 2009;88(10):1828-1833.

12. Devan P, Mahalakshmi N. A study of the performance, emission and combustion characteristics of a compression ignition engine using methyl ester of paradise oil–eucalyptus oil blends. Applied Energy. 2009;86(5):675-680.

13. Imandi SB, Bandaru VR, Somalanka SR, et al. Optimization of medium constituents for the production of citric acid from byproduct glycerol using Doehlert experimental design. Enzyme and Microbial Technology. 2007;40(5):1367-1372.

14. Ravikumar K, Pakshirajan K, Swaminathan T, et al. Optimization of batch process parameters using response surface methodology for dye removal by a novel adsorbent. Chemical Engineering Journal. 2005;105(3):131-138.

15. Box GE, Hunter JS. Multi-factor experimental designs for exploring response surfaces. The Annals of Mathematical Statistics. 1957:195-241.

16. Sarve AN, Varma MN, Sonawane SS. Response surface optimization and artificial neural network modeling of biodiesel production from crude mahua (Madhuca indica) oil under supercritical ethanol conditions using CO 2 as co-solvent. RSC Advances. 2015;5(85):69702-69713.

17. Jena PC, Raheman H, Kumar GP, et al. Biodiesel production from mixture of mahua and simarouba oils with high free fatty acids. Biomass and bioenergy. 2010;34(8):1108-1116.

18. Shanmugaprakash M, Sivakumar V. Development of experimental design approach and ANN-based models for determination of Cr (VI) ions uptake rate from aqueous solution onto the solid biodiesel waste residue. Bioresource technology. 2013;148:550-559.

19. Shanmugaprakash M, Kirthika J, Ragupathy J, et al. Statistical based media optimization and production of naringinase using Aspergillus brasiliensis 1344. International journal of biological macromolecules. 2014;64:443-452.

20. Desai KM, Survase SA, Saudagar PS, et al. Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimization: case study of fermentative production of scleroglucan. Biochemical Engineering Journal. 2008;41(3):266-273.

21. Simon H. Neural networks, a comprehensive foundationPrentice-Hall. Englewood Cliffs, NJ. 1999.

22. Geyikçi F, Kılıç E, Çoruh S, et al. Modelling of lead adsorption from industrial sludge leachate on red mud by using RSM and ANN. Chemical Engineering Journal. 2012;183:53-59.

23. Witek-Krowiak A, Chojnacka K, Podstawczyk D, et al. Application of response surface methodology and artificial neural network methods in modelling and optimization of biosorption process. Bioresource technology. 2014;160:150-160.

24. Moghri M, Shahabadi S, Ismail S, et al. Modeling tensile modulus of (polyamide 6)/nanoclay composites: Response surface method vs. taguchi-optimized artificial neural network. Journal of Vinyl and Additive Technology. 2014.

25. Baskar G, Renganathan S. Optimization of L-asparaginase production by Aspergillus terreus MTCC 1782 using response surface methodology and artificial neural network-linked genetic algorithm. Asia-Pacific Journal of Chemical Engineering. 2012;7(2):212-220.

26. DeCarlo N. The Complete Idiot's Guide to Lean Six Sigma. Penguin; 2007.

27. Özdemir U, Özbay B, Veli S, et al. Modeling adsorption of sodium dodecyl benzene sulfonate (SDBS) onto polyaniline (PANI) by using multi linear regression and artificial neural networks. Chemical Engineering Journal. 2011;178:183-190.

28. Lippmann RP. An introduction to computing with neural nets. ASSP Magazine, IEEE. 1987;4(2):4-22.

29. Ranjan D, Mishra D, Hasan S. Bioadsorption of arsenic: an artificial neural networks and response surface methodological approach. Industrial & Engineering Chemistry Research. 2011;50(17):9852-9863.

30. Ghadge SV, Raheman H. Process optimization for biodiesel production from mahua (Madhuca indica) oil using response surface methodology. Bioresource technology. 2006;97(3):379-384.

31. Venkatesh Prabhu M, Karthikeyan R, Shanmugaprakash M. Modeling and optimization by response surface methodology and neural network–genetic algorithm for decolorization of real textile dye effluent using Pleurotus ostreatus: a comparison study. Desalination and Water Treatment. 2015:1-15.

32. Zafar M, Kumar S, Kumar S, et al. Artificial intelligence based modeling and optimization of poly (3-hydroxybutyrate-co-3-hydroxyvalerate) production process by using Azohydromonas lata MTCC 2311 from cane molasses supplemented with volatile fatty acids: a genetic algorithm paradigm. Bioresource technology. 2012;104:631-641.

33. Giordano PC, Martínez HD, Iglesias AA, et al. Application of response surface methodology and artificial neural networks for optimization of recombinant Oryza sativa non-symbiotic hemoglobin 1 production by Escherichia coli in medium containing byproduct glycerol. Bioresource technology. 2010;101(19):7537-7544.

34. Gen M, Cheng R. Genetic algorithms and engineering optimization. Vol. 7. John Wiley & Sons; 2000.

35. Francis G, Becker K. Biodiesel from Jatropha plantations on degraded land. University of Hohenheim, Germany. 2007.

36. Raheman H, Phadatare A. Diesel engine emissions and performance from blends of karanja methyl ester and diesel. Biomass and Bioenergy. 2004;27(4):393-397.

Figures	
Figure 1	Contour plot showing the combined effect of process variables and FFA (%)
Figure 2	Neural network model with training, validation, test and all prediction set
Figure 3	Comparison of experimental with predicted value obtained by the RSM and ANN model for the prediction of FFA (%) (a) for each experimental run (b) Distribution of residual values
Figure 4	Representative plots generated from the optimization by GA using MATLAB
Figure 5	Chromatogram for biodiesel obtained from simarouba oil (a) Chromatogram for solvent present in simarouba oil sample (b)
Tables	
Table 1	Independent variables and levels used for Box – Behnken design
Table 2	Box – Behnken Design Table for biodiesel production from Simarouba oil
Table 3	ANOVA for response surface quadratic model analysis of variance table
Table 4	Validation data set for experimentally determined, RSM and ANN predicted values of FFA
Table 5	Comparison of predictive abilities of RSM and ANN models
Table 6	Fuel properties of simarouba oil
Table 7	ANN training parameter applied for biodiesel production using *simaruba glauca*

Figure 1

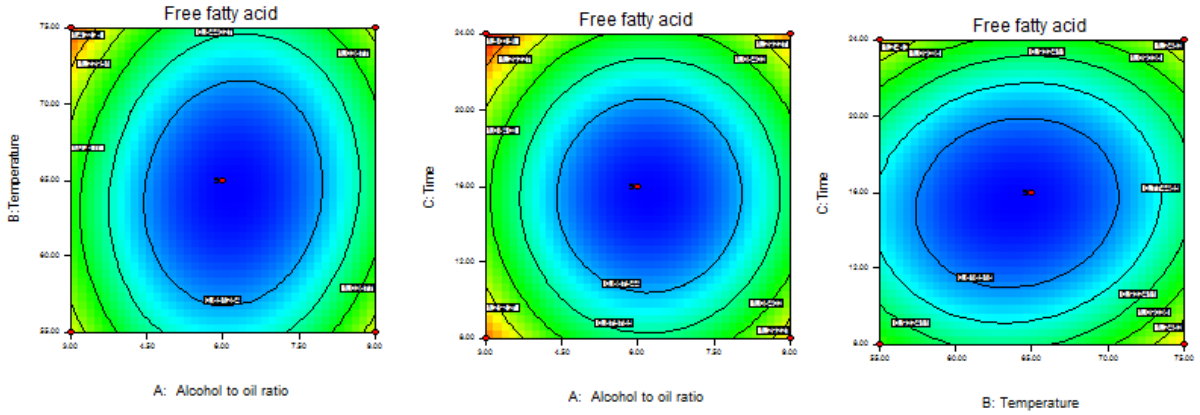

Free fatty acid
B: Temperature
A: Alcohol to oil ratio
Free fatty acid
C: Time
A: Alcohol to oil ratio
Free fatty acid
C: Time
B: Temperature

Figure 2

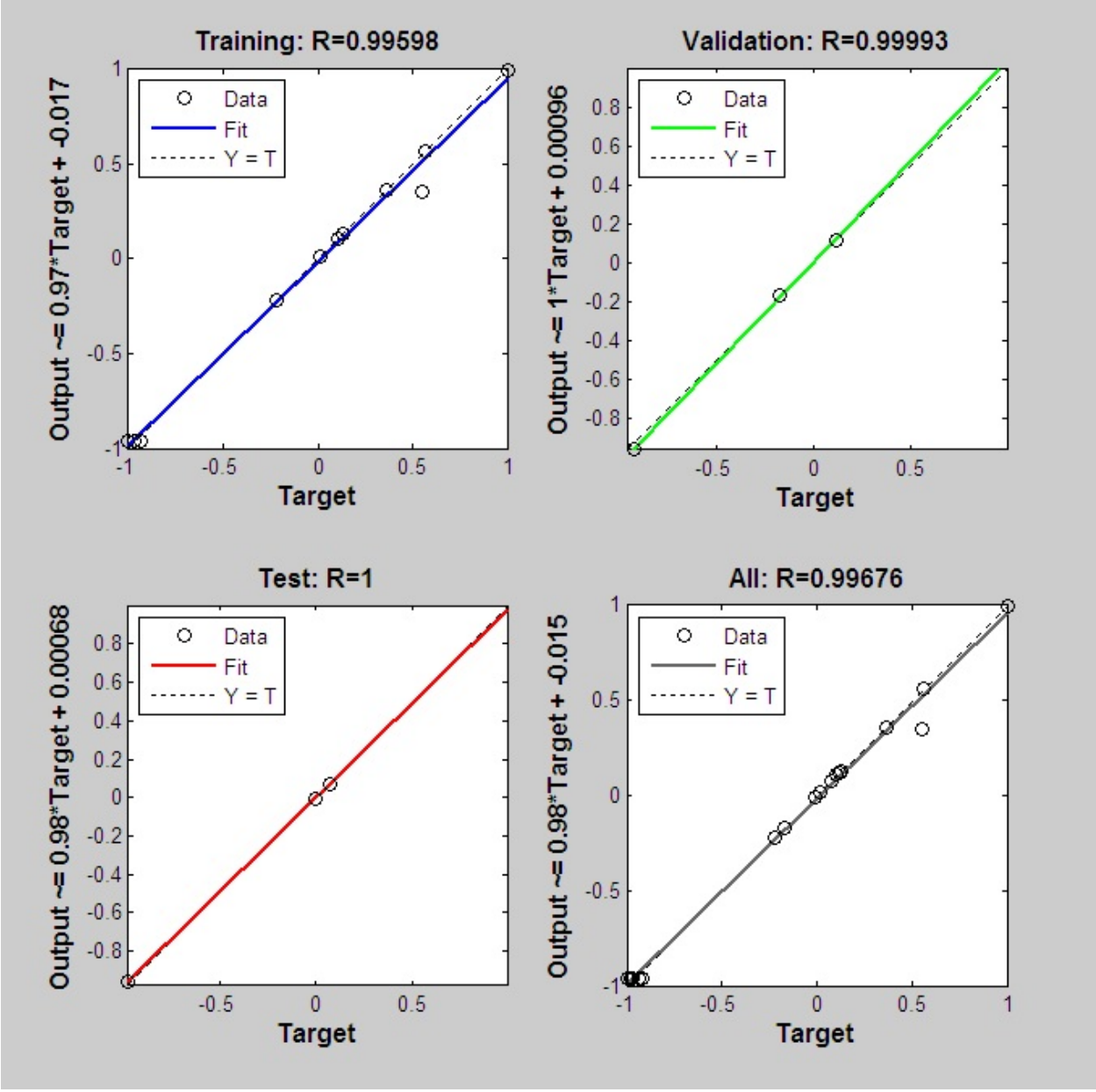

Training: R=0.99598
Output ~= 0.97*Target + -0.017
Target
Data
Fit
Y = T
Validation: R=0.99993
Output ~= 1*Target + 0.00096
Target
Data
Fit
Y = T
Test: R=1
Output ~= 0.98*Target + 0.00068
Target
Data
Fit
Y = T
All: R=0.99676
Output ~= 0.98*Target + -0.015
Target
Data
Fit
Y = T

Figure 3

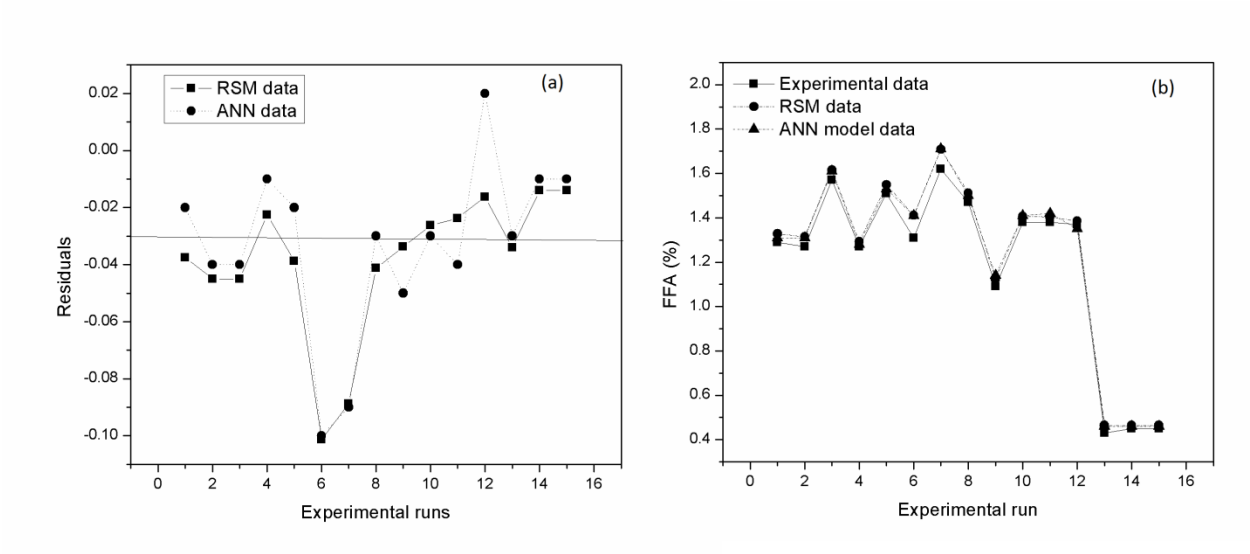


Figure 4

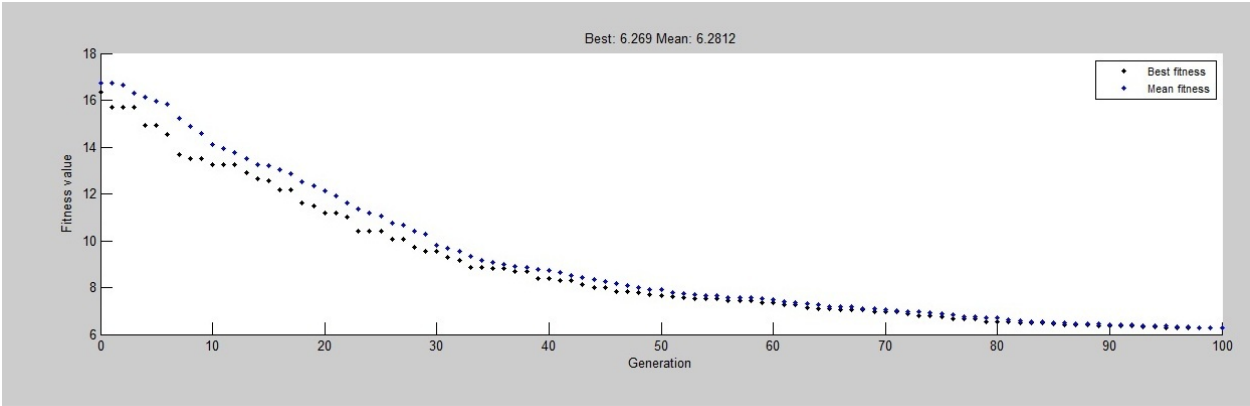


Figure 5

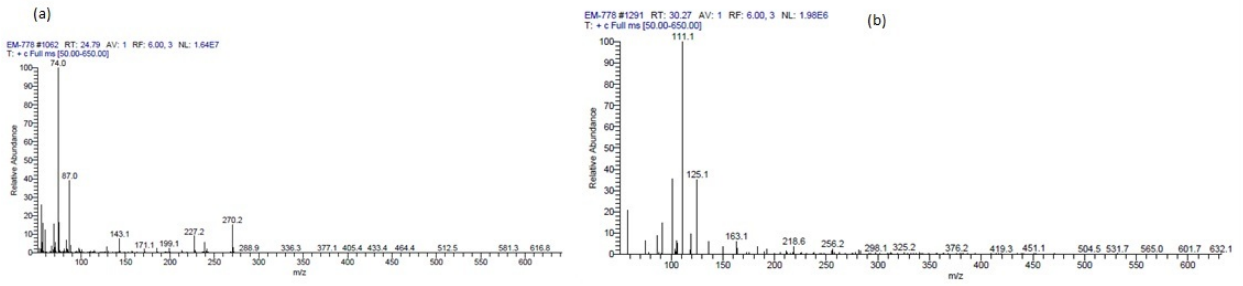


Table 1 Independent variables and levels used for Box – Behnken design

Variables	Symbols	Levels		
		-1	0	1
Oil : Alcohol ratio	M	3	6	9
Feed Temperature (°C)	T	55	65	75
Reaction time (h)	t	8	16	24

Table 2 Box – Behnken Design Table for biodiesel production from Simarouba oil

Run	M (-)	T (°C)	t (h)	FFA (%)		
				Experimental	RSM Predicted	ANN Predicted
1.	3	55	16	1.29	1.3275	1.31
2.	9	55	16	1.27	1.315	1.31
3.	3	75	16	1.57	1.615	1.61
4.	9	75	16	1.27	1.2925	1.28
5.	3	65	8	1.51	1.54875	1.53
6.	9	65	8	1.31	1.41125	1.41
7.	3	65	24	1.62	1.70875	1.71
8.	9	65	24	1.47	1.51125	1.5
9.	6	55	8	1.09	1.12375	1.14
10.	6	75	8	1.38	1.40625	1.41
11.	6	55	24	1.38	1.40375	1.42
12.	6	75	24	1.37	1.38625	1.35
13.	6	65	16	0.43	0.464	0.46
14.	6	65	16	0.45	0.464	0.46
15.	6	65	16	0.45	0.464	0.46

Table 3. ANOVA for response surface quadratic model analysis of variance table

Source	Coefficient	Sum of Squares	df	Mean Square	F value	Prob > F
Model	17.32	3.502849	9	0.389205458	11120.16	< 0.0001
A-M	-0.609	3.289207	9	0.365467418	4228.549	< 0.0001
B-T	0.43	0.056113	1	0.0561125	649.2355	< 0.0001
C-t	-0.18	0.0392	1	0.0392	453.5537	< 0.0001
AB	-0.0025	0.037813	1	0.0378125	437.5	< 0.0001
AC	-0.0006	0.0196	1	0.0196	226.7769	< 0.0001
BC	-0.0009	0.000625	1	0.000625	7.231405	0.0311
A^2	0.059	0.0225	1	0.0225	260.3306	< 0.0001
B^2	0.0036	1.215316	1	1.215316053	14061.51	< 0.0001
C^2	0.007	0.560179	1	0.560179211	6481.412	< 0.0001
Residual		1.020253	1	1.020252895	11804.58	< 0.0001
Lack of Fit		0.000605	7	8.64286E-05		
Pure Error		0.000125	3	4.16667E-05	0.347222	0.7943
Cor Total		3.502849	9	0.389205458	11120.16	< 0.0001

Table 4 Validation data set for experimentally determined, RSM and ANN predicted values of FFA

S.No	M (-)	T (°C)	t (h)	FFA (%)		
				Experimental	RSM Predicted	ANN Predicted
1	3	75	8	1.37	1.30	1.31
2	9	50	3	2.34	2.36	2.34
3	9	60	7	0.67	0.74	0.73
4	6	70	8	0.17	0.19	0.19
5	3	50	8	1.24	1.34	1.32
6	6	75	3	1.65	1.76	1.72

Table 5 Comparison of predictive abilities of RSM and ANN model

Parameters	FFA (%)	
	RSM	ANN
RMSE	0.54	0.29
R^2	0.97	0.99
AAD (%)	7.15	5.91

Table 6. Fuel properties of simarouba oil

Property	**Unit**	**Simarouba oil**	**Simarouba biodiesel**	**Jatropha biodiesel**	**Karanja biodiesel**	**Diesel**
Density	Kg/m^3	960	870	880	876	850
Dynamic viscosity	g/m.s	12.9	8.09	4.2592	8.4096	2.21
Kinematic viscosity	mm^2/s	13.44	9.31	4.84	9.60	2.60
Flash point	°C	142	137	192	187	68
Fire point	°C	147	148	187	177	67
FFA level	%	1.683	0.765	0.24	36.12	0.35

Table 7

ANN training parameter applied for biodiesel production using *simaruba glauca*

Training Parameters used	**Optimal vlaues**
The no. of layers	3
The no. of neurons on the layers	Input:3, Hidden:08, Output:1
Activation functions for hidden and output layer	log-sigmoid transfer function (*logsig*) and linear transfer function(*purelin*)
Training parameter learning rule	Back-Propagation, Levenberg-Marquardt"s learning rule
No. of iteration or epochs	300
Acceptable mean –squared error(MSE) goal	0
Maximum ephos	5000